



UNIVERSIDADE FEDERAL DO ESPÍRITO SANTO  
CENTRO DE CIÊNCIAS DA SAÚDE  
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOTECNOLOGIA

**AMANDA SILVA COUTINHO**

**UTILIZAÇÃO DA BIOINFORMÁTICA NA BUSCA DE NOVOS GENES  
EM OSTEOGÊNESE IMPERFEITA**

VITÓRIA, ES

2018

AMANDA SILVA COUTINHO

**UTILIZAÇÃO DA BIOINFORMÁTICA NA BUSCA DE NOVOS GENES  
EM OSTEOGÊNESE IMPERFEITA**

Dissertação apresentada ao Programa de Pós-Graduação em Biotecnologia do Centro de Ciências da Saúde da Universidade Federal do Espírito Santo, como requisito parcial para obtenção do título de Mestre em Biotecnologia.

Orientadora: Prof<sup>a</sup> Dr<sup>a</sup> Flavia de Paula

VITÓRIA, ES

2018

Dados Internacionais de Catalogação-na-publicação (CIP)  
(Biblioteca Setorial do Centro de Ciências da Saúde da Universidade Federal do Espírito  
Santo, ES, Brasil)

Bibliotecário: Rafael Lima de Carvalho – CRB-6 MG-002926/O

---

Coutinho, Amanda Silva, 1990 -

C871u Utilização da bioinformática na busca de novos genes em  
Osteogênese Imperfeita / Amanda Silva Coutinho – 2018.

67 f. : il.

Orientador: Flavia de Paula.

Dissertação (Mestrado em Biotecnologia) – Universidade Federal do Espírito  
Santo, Centro de Ciências da Saúde.

1. Osteogênese Imperfeita. 2. Biologia Computacional. 3. Genes. I. Paula,  
Flavia de. II. Universidade Federal do Espírito Santo. Centro de Ciências da Saúde.  
III. Título.

CDU: 61

---

AMANDA SILVA COUTINHO

**UTILIZAÇÃO DA BIOINFORMÁTICA NA BUSCA DE NOVOS GENES EM  
OSTEOGÊNESE IMPERFEITA**

Dissertação apresentada ao Programa de Pós-Graduação em Biotecnologia do Centro de Ciências da Saúde da Universidade Federal do Espírito Santo, como requisito parcial para obtenção do título de Mestre em Biotecnologia.

Apresentada em 26 de fevereiro de 2018.

**Prof<sup>a</sup>. Dr<sup>a</sup>. Flávia de Paula**

**Universidade Federal do Espírito Santo**

Orientadora

**Prof<sup>a</sup>. Dr<sup>a</sup>. Flavia Imbroisi Valle Errera**

**Escola Superior de Ciências da Santa Casa de  
Misericórdia de Vitória**

Avaliador interno

**Prof. Dr. Carlos Magno da Costa Maranduba**

**Universidade Federal de Juiz de Fora**

Avaliador externo

**Prof. Dr. Breno Valentim Nogueira**

**Universidade Federal do Espírito Santo**

Avaliador suplente

VITÓRIA, ES

2018

## **AGRADECIMENTOS**

Agradeço a Deus, autor da vida, por me ajudar em cada momento e em cada dificuldade dessa trajetória.

Aos meus pais, por toda dedicação em me educar e me fazer uma pessoa do bem.

Agradeço à minha orientadora, Flavia de Paula, que fez toda a diferença em minha vida quando eu mais precisei, e que foi excelente orientadora. Muito obrigada por toda dedicação e paciência.

À equipe no NGHМ, em especial a Maíra Trancozo, que me ajudou sempre, desde o início até o final desse projeto, colaborando com seus conhecimentos e sua amizade.

A Elaine Stur, por me ensinar a utilizar ferramentas importantes para este trabalho e às amigas de turma do mestrado Dalila e Jucimara que sempre me incentivaram.

Agradeço ao meu esposo Paulo, por me ajudar a nunca desistir dos meus objetivos, pelo apoio e incentivo.

Agradeço à UFES e ao PPG Biotecnologia pela oportunidade de desenvolver a pesquisa e à FAPES, MCTI, CNPQ, MEC e CAPES pelo apoio financeiro.

## RESUMO

COUTINHO, A.S. **Utilização da bioinformática na busca de novos genes em osteogênese imperfeita**. 2018. 67f. Dissertação de Mestrado em Biotecnologia – Programa de Pós-Graduação em Biotecnologia, UFES, Espírito Santo. Brasil.

A osteogênese imperfeita (OI) é uma doença genética rara do tecido conjuntivo, causada por mutações em genes que participam, em geral, da formação óssea. A maioria dos pacientes é portadora de mutações nos genes que codificam o colágeno tipo 1, mas já foram descritas mutações em mais de 17 outros genes causando OI e ainda existe uma busca constante de novos genes na área científica. Entre as estratégias de diagnóstico molecular destaca-se a técnica de sequenciamento de nova geração (NGS), que pode sequenciar vários genes presentes em uma plataforma customizada, gerando uma grande quantidade de dados genômicos. Esses dados se tornam preciosas fontes de informação na busca de novos genes relacionados a doenças. O objetivo desta pesquisa foi realizar a busca de novos genes potencialmente causadores de OI por meio de recursos de bioinformática. Foram utilizadas estratégias de filtragem pelo programa Microsoft Office Excel 2013, bem como análises de predição de mutação. Como referência genômica foram utilizados os bancos de dados *Ensembl* e *National Center for Biotechnology Information*. Foram selecionados quatro pacientes diagnosticados clinicamente com OI que foram submetidos à técnica de NGS e apresentaram resultados normais para os genes conhecidos. Com o intuito de selecionar uma lista de genes candidatos na plataforma customizada de NGS que estivessem relacionados com os sintomas de OI, foi realizada uma busca de genes no banco de dados *Ensembl* envolvidos com as vias metabólicas de formação óssea, cartilaginosa ou de colágeno, que identificou 643 genes. A lista de genes candidatos foi comparada com os genes sequenciados dos pacientes, onde foram selecionados 70 genes em comum para análise. Foram realizadas filtrações *in silico* de forma a selecionar alterações raras na população, preditas como patogênicas e que efetivamente codifiquem uma proteína ou uma molécula de RNA funcional. Os resultados mostraram que o paciente P.1 é portador de uma mutação em heterozigose potencialmente patogênica no gene *ALX1*. O paciente P.2 apresentou apenas uma alteração no gene *COL6A3* que foi predita como

polimorfismo. O paciente P.3 apresentou mutações patogênicas em heterozigose nos genes *ALPL* e *FKBP10*. No paciente P.4 foram encontradas mutações patogênicas em heterozigose nos genes *P3H1* e *RYR1*. Entre os cinco genes identificados, sabe-se que dois deles, *FKBP10* e *P3H1*, estão relacionados com a OI de herança autossômica recessiva. Também já é descrito que mutações no gene *ALPL* causam sintomas clínicos semelhantes a OI, podendo confundir o diagnóstico. Assim, o presente estudo identificou dois genes, *ALX1* e *RYR1*, potencialmente causadores de OI. O gene *ALX1* tem um papel importante no desenvolvimento craniano e dos membros, pois atua na formação da cartilagem. Já o *RYR1* codifica a rianodina, um importante receptor de cálcio nos osteoblastos. Estudos funcionais dos genes identificados são necessários para validar esta hipótese em pesquisas futuras. Os resultados deste trabalho sugerem que ferramentas de bioinformática podem direcionar a busca por novos genes relacionados a doenças genéticas. A caracterização de novas mutações em genes relacionados com OI auxilia no planejamento de estratégias mais eficientes que permitam o diagnóstico molecular da doença e o aconselhamento genético.

Palavras-chave: Next Generation Sequencing. Bioinformática. Osteogênese Imperfeita. Gene *ALX1*. Gene *RYR1*.

# USE OF BIOINFORMATICS IN THE SEARCH FOR NEW GENES IN OSTEOGENESIS IMPERFECTA

## ABSTRACT

COUTINHO, A.S. **Use of bioinformatics in the search for new genes in osteogenesis imperfecta**. 2018. 67f. Master's Dissertation in Biotechnology - Post-Graduate Program in Biotechnology, UFES, Espírito Santo. Brazil.

Osteogenesis imperfecta (OI) is a rare genetic disease of connective tissue caused by mutations in genes that generally participate in bone formation. Most patients carry mutations in the genes encoding type 1 collagen, but mutations in more than 17 other genes causing OI have been described and there is still a constant search for new genes in the scientific field. Among the molecular diagnostic strategies, the new generation sequencing method (NGS) stands out, which can sequence several genes present in a customized platform generating a large amount of genomic data. These data become precious sources of information in the search for new genes related to diseases. The objective of this research was to perform the search for new genes potentially causing OI through bioinformatics resources. We used filtering strategies by the Microsoft Office Excel 2013 program, as well as mutation prediction analyzes. As a genomic reference, the Ensembl and National Center for Biotechnology Information databases were used. We selected four patients diagnosed clinically with OI who were submitted to the NGS technique and presented normal results for the known genes. In order to select a list of candidate genes in the NGS custom platform that were related to OI symptoms, a search of genes in the Ensembl database involved with the metabolic pathways of bone, cartilage or collagen formation was performed, which identified 643 genes. The list of candidate genes was compared to the sequenced genes of the patients, where 70 genes in common were selected for analysis. *In silico*, filtrations were performed in order to select rare changes in the population, predicted as pathogenic and that effectively encode a functional RNA protein or molecule. The results showed that patient P.1 carries a potentially pathogenic heterozygous mutation in the *ALX1* gene. Patient P.2 presented only one



alteration in the *COL6A3* gene that was predicted as polymorphism. Patient P.3 presented pathogenic mutations in heterozygosity in the *ALPL* and *FKBP10* genes. In patient P.4, pathogenic mutations in heterozygosity were found in the *P3H1* and *RYR1* genes. Among the five genes identified, two of them, *FKBP10* and *P3H1*, are known to be related to autosomal recessive OI. It has also been described that mutations in the *ALPL* gene cause clinical symptoms similar to OI, which may confuse the diagnosis. Thus, the present study identified two genes, *ALX1* and *RYR1*, potentially causing OI. The *ALX1* gene plays an important role in cranial and limb development, as it acts on the formation of cartilage. *RYR1* encodes ryanodine, an important calcium receptor in osteoblasts. Functional studies of the identified genes are necessary to validate this hypothesis in future research. The results of this work suggest that bioinformatics tools may direct the search for new genes related to genetic diseases. The characterization of new mutations in OI-related genes enables the planning of more efficient strategies that allow molecular diagnosis of the disease and genetic counseling.

Key words: Next Generation Sequencing. Bioinformatics. Osteogenesis imperfecta. *ALX1* gene. *RYR1* gene.

## LISTA DE FIGURAS

Figura 1: Figura representativa do exemplar dissecado por Willem Vrolik (BALJET, 2002).....	21
Figura 2: Radiografia do exemplar encontrado pelo professor John Garstang (GRAY, 1969).....	22
Figura 3: Heredogramas dos pacientes P.1 (A), P.2 (B), P.3 (C) e P.4 (D). ....	31
Figura 4: Fluxograma da metodologia. ....	33
Figura 5: ALX1 sequência selvagem. ....	46
Figura 6: ALX1 Sequência mutada.....	46

## LISTA DE TABELAS

Tabela 1: Classificação de Sillence e colaboradores (1979). .....	24
Tabela 2: Classificação genética da osteogênese imperfeita. ....	24
Tabela 3: Caracterização dos pacientes estudados. ....	30
Tabela 4: Alterações genéticas encontradas no paciente 1.....	41
Tabela 5 - Alterações genéticas encontradas no paciente 2. ....	42
Tabela 6 - Alterações genéticas encontradas no paciente 3. ....	43
Tabela 7 - Alterações genéticas encontradas no paciente 4 .....	44

## LISTA DE ABREVIATURAS E SIGLAS

AC: Contagem de alelos (do inglês, *Allele count*)

ACMG: Colégio Americano de Genética Médica e Genômica (do inglês, *American College of Medical Genetics and Genomics*)

ALPL: Fosfatase alcalina, fígado / osso / rim (do inglês, *alkaline phosphatase, liver/bone/kidney*)

ALX1: Gene *ALX* homeobox 1

ALX3: Gene *ALX* homeobox 3

ALX4: Gene *ALX* homeobox 4

APS: Persulfato de amônio

AVSIFT: do inglês, *average SIFT*

AXIN2: Gene *AXIN2* (nome completo oficial)

BMP1: Gene proteína morfogenética óssea 1 (do inglês, *bone morphogenetic protein 1*)

BWA: Alinhador de Burrows-Wheeler (do inglês, *Burrows-Wheeler Aligner*)

ClinVar: Variação clínica (do inglês, *Clinical Variation*)

CART-1: Gene transcrição regulada por cocaína e anfetaminas (do inglês, *cocaine-and amphetamine-regulated transcript 1*)

COL1A1: Gene do colágeno tipo I, cadeia alfa 1 (do inglês, *collagen type I alpha 1 chain*)

COL1A2: Gene do colágeno tipo I, cadeia alfa 2 (do inglês, *collagen type I alpha 2 chain*)

COL5A2: Gene do colágeno tipo V, cadeia alfa 2 (do inglês, *collagen type V alpha 2 chain*)

COL6A1: Gene do colágeno tipo VI, cadeia alfa 1 (do inglês, *collagen type VI alpha 1 chain*)

COL6A2: Gene do colágeno tipo VI, cadeia alfa 2 (do inglês, *collagen type VI alpha 2 chain*)

COL6A3: Gene do colágeno tipo VI, cadeia alfa 3 (do inglês, *collagen type VI alpha 3 chain*)

COMP: Gene da proteína da matriz oligomérica da cartilagem (do inglês, *cartilage oligomeric matrix protein*)

CREB3L1: Gene da proteína 3 de ligação ao elemento responsivo ao cAMP como 1 (do inglês, *cAMP responsive element binding protein 3 like 1*).

CRTAP: Gene da proteína associada à cartilagem (do inglês, *cartilage associated protein*)

DNA: Ácido desoxirribonucleico

FKBP10: Gene da proteína de ligação FK506 10 (do inglês, *FK506 binding protein 10*)

Func.refGene: Função do gene

EDTA: Ácido etilenodiaminotetracético

EGL: Laboratório de Genética Molecular Emory (do inglês, *Emory Molecular Genetics Laboratory*)

ES: Espírito Santo

EVC: Subconjunto complexo ciliar de EVC 1 (do inglês, *EVC ciliary complex subunit 1*)

ExAC: Consórcio de agregação de exoma (do inglês, *Exome Aggregation Consortium*)

FND: Displasia frontonasal (do inglês, *Frontonasal dysplasia*)

GATK: Kit de ferramentas de análise do genoma (do inglês, *Genome Analysis Toolkit*)

GDF5: Gene do fator de diferenciação do crescimento 5 (do inglês, *growth differentiation factor 5*)

GeneDx: Critérios Gerais de Afirmação da Classificação Variante (do inglês, *General Variant Classification Assertion Criteria*)

GNAS: Gene Locus complexo GNAS (do inglês, *GNAS complex locus*)

HINSG: Hospital Infantil Nossa Senhora da Glória

HSP47: Proteína de choque térmico 47 (do inglês, *Heat-Shock Protein 47*)

ICSL: Laboratório de Serviços Clínicos Illumina (do inglês, *Illumina Clinical Services Laboratory*)

IVC: Classificação Variante Invitae (do inglês, *Invitae Variant Classification*)

IFITM5: Gene da proteína transmembrana induzida por interferon 5 (do inglês, *interferon induced transmembrane protein 5*)

kDa: Kilodalton

P3H1: Gene da prolil-3 hidroxilase 1

LH2: Gene do procolágeno-lisina, 2-oxoglutarato 5-dioxigenase 2

MBTPS2: Gene do fator de transcrição ligado à membrana peptidase, sítio 2 (do inglês, *membrane bound transcription factor peptidase, site 2*)

MH: Hipertermia maligna

MIM: Herança Mendeliana no Homem (do inglês, *Mendelian Inheritance in Man*)

Min: minutos

mL: mililitros

mM: milimolar

NA: Não avaliado

NCBI: Centro Nacional de Informação Biotecnológica (do inglês, *National Center for Biotechnology Information*)

NGS: Sequenciamento de próxima geração (do inglês, *Next Generation Sequencing*)

OI: Osteogênese Imperfeita

*P3H1*: Gene da prolil-3 hidroxilase 1

PCR: Reação em Cadeia da Polimerase (do inglês, *Polymerase chain reaction*)

PEDF: Fator de pigmento derivado do epitélio (do inglês, *Pigment Epithelium-Derived Factor*)

PPIase: Peptidyl-prolyl cis–trans isomerase

*PPIB*: Gene da peptidilprolyl isomerase B

*PLOD2*: Gene do procólágeno-lisina, 2-oxoglutarato 5-dioxigenase 2

Q-PCR: PCR quantitativo (do inglês, *quantitative PCR*)

rpm: Rotações por minuto

*RYR1*: Gene receptor de ryanodina 1 (do inglês, *ryanodine receptor 1*)

SDS: Dodecil sulfato de sódio

Seg: segundos

*SERPINH1*: Gene H da família Serpin 1 (do inglês, *serpin family H member 1*)

*SERPINF1*: Gene F da família Serpin 1 (do inglês, *serpin family F member 1*)

SNP: Polimorfismo de nucleotídeo único (do inglês, *Single-nucleotide polymorphism*)

SNV: Variantes de nucleotídeos únicos (do inglês, *single nucleotide variants*)

SP7/Osterix: Gene do fator de transcrição SP7 (do inglês, *SP7 Transcription Factor*)

SPARC: Gene da proteína secretada ácida e rica em cisteína (do inglês, *secreted protein acidic and cysteine rich*)

TBE: Tampão tris/borato/EDTA

TEMED: Tetrametiletenodiamina (do inglês, *Tetramethylethylenediamine*)

TMEM38B: Gene da proteína transmembrana 38B (do inglês, *Transmembrane Protein 38B*)

TNSALP: Gene da fosfatase alcalina, fígado / osso / rim (do inglês, *alkaline phosphatase, liver/bone/kidney*)

TRIC-B: Canal de cátion intracelular trimérico tipo B (do inglês, *Trimeric Intracellular Cation Channel Type B*)

TRIP11: Gene Interagente do receptor de hormônio da tireóide 11 (do inglês, *thyroid hormone receptor interactor 11*)

UFES: Universidade Federal do Espírito Santo

USH2A: Gene da Síndrome de Usher tipo IIa (do inglês, *Usher syndrome type IIa*)

WNT1: Gene da família do sítio de integração MMTV sem alças, membro 1 (do inglês, *wingless-type MMTV integration site family, member 1*)

µL: microlitro

°C: graus Celsius



## SUMÁRIO

1. INTRODUÇÃO .....	18
2. REVISÃO BIBLIOGRÁFICA .....	20
2.1. HISTÓRICO DA DOENÇA .....	20
2.2. CARACTERIZAÇÃO DA DOENÇA OSTEOGÊNESE IMPERFEITA .....	23
2.2.1. Manifestações Clínicas .....	23
2.2.2. Classificação .....	23
2.2.3. Genes .....	26
2.3. NEXT GENERATION SEQUENCING .....	27
3. OBJETIVOS .....	28
3.1. OBJETIVO GERAL .....	28
3.2. OBJETIVOS ESPECÍFICOS .....	28
4. METODOLOGIA .....	29
4.1. AMOSTRA .....	29
4.2. SELEÇÃO DOS PACIENTES ESTUDADOS .....	32
4.3. MÉTODOS .....	32
4.4. ANÁLISE EXPERIMENTAL .....	34
4.4.1. Next Generation Sequencing .....	34
4.5. ANÁLISE DE BIOINFORMÁTICA .....	35
4.5.1. Parâmetros de bioinformática utilizados na seleção das variantes sequenciadas .....	35
4.5.2. Etapas de filtragem dos dados .....	36
4.5.3. Seleção de genes candidatos .....	37
4.5.4. Exclusão das mutações sinônimas .....	38
4.5.5. Análise de predição de mutações e proteínas .....	38
5. RESULTADOS .....	41
6. DISCUSSÃO .....	47
6.1. VARIAÇÕES GENÉTICAS IDENTIFICADAS .....	47
6.1.1. Paciente P.1 .....	47
6.1.1.1. Variante no gene <i>ALX1</i> .....	47
6.1.1.2 Variantes no gene <i>COL6A1</i> .....	47
6.1.1.3. Variante no gene <i>COL6A2</i> .....	48

6.1.1.4 Conclusão sobre as análises das variantes do Paciente P.1 .....	48
6.1.2. Paciente P.2 .....	49
6.1.3. Paciente P.3 .....	49
6.1.3.1. Variantes no gene <i>COL6A3</i> .....	50
6.1.3.2. Variantes no gene <i>AXIN2</i> .....	50
6.1.3.3. Variantes no gene <i>ALPL</i> .....	51
6.1.3.4. Conclusão sobre as análises das variantes do Paciente P.3 .....	52
6.1.4. Paciente P.4 .....	52
6.1.4.1. Variantes no gene <i>AXIN2</i> .....	52
6.1.4.2. Variantes no gene <i>COL5A2</i> .....	53
6.1.4.3. Variantes no gene <i>TRIP11</i> .....	53
6.1.4.4. Variante no gene <i>RYR1</i> .....	53
6.1.4.5. Conclusão sobre as análises das variantes do Paciente P.4 .....	54
6.2. GENES POTENCIALMENTE PATOGÊNICOS.....	54
6.2.1. Gene <i>ALX1</i> .....	54
6.2.2. Gene <i>RYR1</i> .....	56
7. CONSIDERAÇÕES ADICIONAIS .....	58
8. CONCLUSÕES.....	58
9. REFERÊNCIAS.....	59
ANEXO I – AMOSTRA DA PLANILHA DO NGS COM OS PRINCIPAIS PARÂMETROS ANALISADOS. ....	64

## 1. INTRODUÇÃO

A osteogênese imperfeita (OI) é um grupo heterogêneo de doenças genéticas raras do tecido conjuntivo (HUNG; TAN; HIEN, 2017; KANG; ARYAL A.C.; MARINI, 2017). É também conhecida como a doença dos ossos frágeis ou ossos de vidro. Essa síndrome envolve uma alteração no colágeno tipo 1, que é o componente orgânico do osso (MARINI et al., 2017). No entanto, com estudos recentes foi comprovado que muitos pacientes com OI não mostraram mudanças nos genes que codificam a produção de colágeno. Hoje o conceito foi ampliado, sendo a osteogênese imperfeita definida como uma síndrome causada por uma mudança genética com níveis de complexidade variáveis (MARINI et al., 2017).

O diagnóstico é fornecido pelo exame clínico e testes radiológicos, e pode ser feito ainda na gravidez, dependendo do tipo de OI, por meio de exames de ultrassonografia, testes laboratoriais e pelo histórico familiar (BYERS et al., 2006). A classificação de Sillence (1979), tem sido empregada para agrupar esses pacientes de acordo com as características clínicas e radiológicas (HUNG; TAN; HIEN, 2017).

Manifestações clínicas típicas são fragilidade óssea, baixa massa óssea, deformidades no esqueleto e baixa estatura. É também uma desordem sistêmica que envolve estruturas extra esqueléticas. Isso engloba várias manifestações clínicas incluindo dentinogênese imperfeita, esclerótica azul-cinza, deficiência auditiva, hipermobilidade articular, hipotonia muscular, doença pulmonar restritiva e anormalidades cardiovasculares (GONÇALVES; MEYER; SATO, 2017; KANG; ARYAL A.C.; MARINI, 2017). Em geral, a prevalência de OI é estimada em 1 a cada 15.000 – 20.000 nascimentos. A severidade de OI varia de letalidade perinatal e deformidade severa até tipos de formas muito suaves sem deformidade. A heterogeneidade genética de OI é mais complicada pela extensa variabilidade fenotípica de cada locus genético e pelos diferentes modos de herança (KANG; ARYAL; MARINI, 2017).

Pelo fato de a OI ser uma doença que pode ser causada por mutações em diferentes genes, o diagnóstico conclusivo se dá através de diagnóstico molecular do paciente em busca de mutações nos principais genes envolvidos com a formação óssea.

Entretanto, existem pacientes que possuem os sintomas da doença, mas não têm a causa genética conhecida, pois não possuem mutações em nenhum dos genes já associados à OI.

Apesar de incurável, faz-se necessário o diagnóstico precoce da OI, como forma de ampliar as formas de tratamento, prevenir novas fraturas e suas complicações, possibilitando o manejo adequado e com melhor prognóstico e qualidade de vida (GONÇALVES; MEYER; SATO, 2017). Identificar a causa molecular auxilia no diagnóstico e ajuda a encontrar estratégias para o tratamento da doença.

Uma das formas de se fazer o diagnóstico molecular é por meio do sequenciamento de nova geração (NGS). Essa técnica gera uma grande quantidade de dados, que muitas vezes não são explorados, pois, em geral, os pesquisadores procuram por genes já conhecidos como forma de diagnosticar a doença. Todavia, podem existir outros genes que, quando alterados, causem a doença ou alterem a manifestação clínica de um paciente, que estejam relacionados à formação óssea, cartilaginosa ou de colágeno e que não foram ainda pesquisados.

Utilizando ferramentas computacionais e *softwares* que possuem mecanismos de filtragem de dados, é possível selecionar, nos dados gerados pelo sequenciamento, genes alterados em pacientes, e investigar se as alterações encontradas têm alguma relação com a doença que ele manifesta.

Frente à necessidade de contribuir na divulgação do conhecimento dessa enfermidade entre a comunidade científica, esse estudo tem por objetivo buscar novos genes associados à OI, utilizando ferramentas de bioinformática a partir do sequenciamento de DNA de nova geração de pacientes com OI. Dessa forma, a análise destes dados pode trazer informações ricas sobre alterações genéticas relacionadas à doença ou sobre as vias metabólicas envolvidas com ela envolvidas.

## 2. REVISÃO BIBLIOGRÁFICA

### 2.1. HISTÓRICO DA DOENÇA

A partir do final do século XVII, os relatos de condições associadas a ossos anormalmente frágeis começaram a aparecer. Em 1788, Olof Jakob Ekman, um cirurgião militar sueco, descreveu uma família em que quatro gerações sofreram de uma condição hereditária de fragilidade óssea que ele chamava de *osteomalacia congenita* (WEIL, 1981). Algumas décadas mais tarde, o patologista francês Lobstein relatou em seu *Traité d'anatomie pathologique générale* (1833) sobre três crianças de uma família afetada por *fragilitas ossium*, que ele denominou como *osteopsathyrosis idiopathica*. Notavelmente ele observou também que mais órgãos eram afetados pela doença, sem, contudo, dar maiores detalhes. Em 1849, o professor de Anatomia, Fisiologia, História Natural e Cirurgia Teórica da *Athenaeum Illustre of Amsterdam*, Wilhem Vrolik, descreveu uma criança com encurtamento e extremidades deformadas, que morreu três dias após o nascimento: “No esqueleto é manifestada lesão por uma multidão de fraturas. Nenhuma costela nem ossos longos foram poupados: vários, na verdade, até mesmo exibem duas ou três fraturas. Sem dúvida esta condição singular do esqueleto é atribuível à ossificação imperfeita” (BALJET, 2002). Willem Vrolik foi um dos primeiros de sua época a perceber que a maioria das displasias ósseas não era adquirida após o nascimento, mas poderia ser uma insuficiência na “energia generativa” durante gravidez. Essa afirmativa foi confirmada pelo anatomista quando, estudando peças anatômicas, constatou o comprometimento da ossificação primária, e não uma degeneração secundária. O caso do recém-nascido com inúmeras fraturas falecido três dias após o nascimento (Figura 1) foi descrito em seu livro *Patologias Anatômicas*, e decorreu da análise de um exemplar doado à coleção de seu pai, o também anatomista Gerard Vrolik (BALJET, 2002).

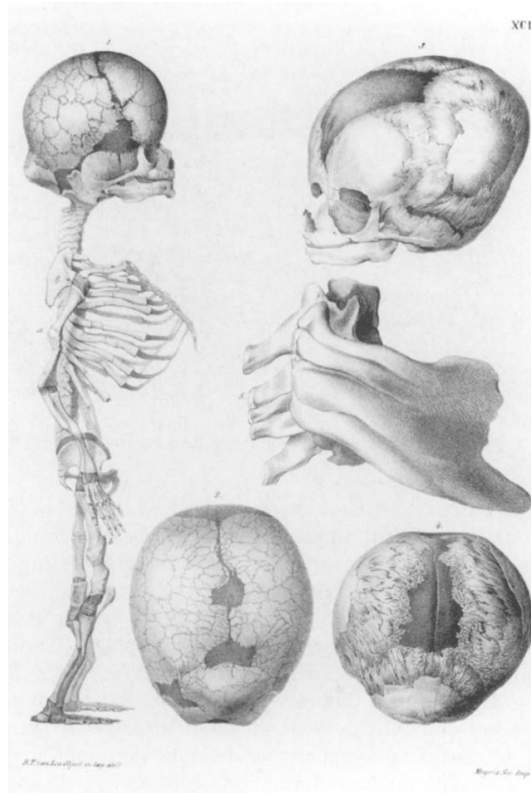


Figura 1: Figura representativa do exemplar dissecado por Willem Vrolik (BALJET, 2002).

Ao dissecar o exemplar, o professor Vrolik constatou que o recém-nascido apresentava deficiências no desenvolvimento esquelético, que os ossos do crânio eram unidos por bordas arredondadas, além de apresentarem fraturas e irregularidades, que a face era curta e estreita, e que os demais ossos estavam fraturados em diversos pontos, sendo que, em algumas das fraturas, foram identificados calos ósseos decorrentes de prévia regeneração intrauterina. Um dos desenhos elaborados pelo anatomista para descrever esse exemplar foi intitulado, em latim, *Osteogenesis Imperfecta*, sendo essa a primeira vez que esse termo foi utilizado na literatura (BALJET, 2002; BALJET; GROEN, 1984).

Assim, em meados do século XIX, uma doença chamada *osteogenesis imperfecta* congênita foi definida como uma condição inerentemente diferente do raquitismo ou da “doença inglesa” pós-natal adquirida que se espalhou ao longo do século da revolução industrial (O’RIORDAN; BIJVOET, 2014).

Por cerca de meio século, *osteopsathyrosis idiopathica* e *osteogenesis imperfecta* congênita permaneceram duas doenças diferentes, até que o patologista suíço Emil Looser forneceu evidências histológicas de que ambas as enfermidades eram manifestações do mesmo transtorno. Ele reconheceu que a OI era uma condição sistêmica rara de etiologia desconhecida, caracterizada por “desenvolvimento

imperfeito de osso” e fez a primeira classificação em dois tipos, de acordo com o momento em que ocorreu a primeira fratura: congênita (tipo Vrolik), com fraturas já no período perinatal, e tardia (tipo Ekman-Lobstein), quando as fraturas ocorreram após o primeiro ano de vida. Essa primeira classificação de OI foi posteriormente modificada por Seedorf, que subclassificou a OI tardia em *OI tardia gravis*, quando as primeiras fraturas ocorreram no primeiro ano de vida devido ao desenvolvimento subsequente de deformidades esqueléticas graves e *OI tardia levis*, quando as fraturas ocorreram mais tarde (WEIL, 1981).

Em 1969, Gray realizou a primeira descrição de um caso de Osteogênese Imperfeita identificado em uma múmia egípcia da 21ª dinastia, datada de aproximadamente 1.000 anos antes de Cristo, encontrada às margens do rio Nilo pelo professor John Garstang (GARSTANG, 1907 apud GRAY, 1969). Por meio de exames radiológicos e de observações realizadas pelo professor Gray, foi constatado que o exemplar apresentava dentinogênese imperfeita, ossos longos deformados e ossos dos membros inferiores arqueados (Figura 2), e que o córtex era composto de linhas esponjosas e onduladas, características típicas da doença (GRAY, 1969; LOWENSTEIN, 2009).



Figura 2: Radiografia do exemplar encontrado pelo professor John Garstang (GRAY, 1969).

Esse achado demonstrou que a doença já acometia povos da antiguidade, não sendo, dessa forma, uma condição recente, além de oferecer alguns subsídios históricos para avaliação da evolução da doença (LOWENSTEIN, 2009).

## 2.2. CARACTERIZAÇÃO DA DOENÇA OSTEOGÊNESE IMPERFEITA

### 2.2.1. Manifestações Clínicas

Osteogênese Imperfeita (OI) é o termo coletivo para um grupo heterogêneo de síndromes do tecido conjuntivo, caracterizados principalmente por fraturas ósseas ao longo da vida. As características principais dessa doença envolvem fragilidade óssea, predispondo a fraturas recorrentes e deformidades esqueléticas. As características secundárias incluem esclerótica azul, perda de audição, dentinogênese imperfeita, escoliose, comprometimento da função pulmonar, anormalidades da válvula cardíaca, fraqueza muscular e hipermobilidade das articulações (JIA et al., 2017).

### 2.2.2. Classificação

Muitas formas novas de OI ou doença dos ossos frágeis foram descobertas na última década. A nomenclatura tradicional, de acordo com Sillence et al. (1979), define os tipos de OI de acordo com a gravidade da doença. As formas recentemente descobertas e especialmente o crescimento do conhecimento sobre a doença têm desafiado essa classificação. A nomenclatura de Sillence et al. (1979) foi substancialmente ampliada com uma mudança gradual de classificação fenotípica a genotípica. Desde a primeira descrição de OI, em 1788 (BALJET, 2002), a nomenclatura e a classificação da OI evoluíram substancialmente.

Em 1979, Sillence e colaboradores observaram uma grande variabilidade clínica entre seus pacientes e com isso propuseram, em suas pesquisas, uma classificação para a Osteogênese Imperfeita.

Essa classificação foi feita por meio de dados clínicos e radiológicos, que permitiram categorizar a doença em quatro tipos que compreendem do mais leve até o inviável à vida (Tabela 1)



Tabela 1: Classificação de Silience e colaboradores (1979).

Classificação	Níveis de gravidade das características clínicas
Tipo I	Leve
Tipo II	Letal pré e pós nascimento
Tipo III	Grave
Tipo IV	Moderada

Atualmente, com as descobertas de novos genes de herança recessiva, foram feitas reclassificações que se baseiam nos dados histológicos e genéticos. A classificação mais recente, feita por Marini et al. (2017), está demonstrada na tabela 2, que atribui à OI os tipos I até o XVIII, sendo que alguns casos ainda continuam sem classificação.

Tabela 2: Classificação genética da osteogênese imperfeita.

Gene mutado	Proteína codificada	Tipo de Osteogênese imperfeita	Herança	Características clínicas
<i>COL1A1</i> ou <i>COL1A2</i>	Colágeno $\alpha 1(I)$ ( <i>COL1A1</i> ) ou $\alpha 2(I)$ ( <i>COL1A2</i> )	I, II, III ou IV*	Autossômica Dominante	Fenótipo clássico
<i>IFITM5</i>	Proteína semelhante a proteína transmembrana induzida por interferon restrita a ossos (BRIL; também conhecido como IFM5)	V	Autossômica Dominante	Deformidade esquelética normal a severa, ossificações na membrana intraóssea, banda radiodensa da cabeça radial, esclera normal a azul e, às vezes, perda auditiva
<i>SERPINF1</i>	Fator de pigmento derivado do epitélio (PEDF)	VI	Autossômica Recessiva	Deformidade esquelética moderada a severa, presença de osteóides, padrão de osso lamelar semelhante a escamas de peixe, com início na infância
<i>CRTAP</i>	Proteína associada a cartilagem ( <i>CRTAP</i> )	VII	Autossômica Recessiva	Rizomelia severa com esclera branca
<i>P3H1</i>	Prolyl 3-hydroxylase 1 ( <i>P3H1</i> )	VIII	Autossômica Recessiva	Rizomelia severa com esclera branca

<i>PPIB</i>	Peptidyl-prolyl <i>cis-trans</i> isomerase B (PPIase B)	IX	Autossômica Recessiva	Deformidade óssea severa com esclera cinzenta
<i>SERPINH1</i>	Serpin H1 (também conhecida como HSP47)	X	Autossômica Recessiva	Deformidade esquelética severa, esclera azul, dentinogênese imperfeita, anormalidades na pele e hérnia inguinal
<i>FKBP10</i>	65 kDa FK506- proteína de ligação (FKBP65)	XI	Autossômica Recessiva	Deformidade esquelética leve a severa, esclera normal a cinzenta e contraturas congênitas
<i>PLOD2</i>	Lysyl hydroxylase 2 (LH2)	Sem tipo	Autossômica Recessiva	Deformidades esqueléticas moderadas a severas e contraturas progressivas das juntas
<i>BMP1</i>	Proteína morfogenética do osso 1 ( <i>BMP1</i> )	XII	Autossômica Recessiva	Deformidade esquelética leve a severa e hérnia umbilical
<i>SP7</i>	Fator de transcrição <i>SP7</i> (também conhecido como osterix)	XIII	Autossômica Recessiva	Deformidade esquelética severa com atraso no aparecimento dos dentes e hipoplasia facial
<i>TMEM38B</i>	Canal de cátion intracelular trimérico tipo B (TRIC-B; também conhecido como TM38B)	XIV	Autossômica Recessiva	Deformidade óssea severa com esclera normal a azul
<i>WNT1</i>	Proto-oncogene Wnt-1 ( <i>WNT1</i> )	XV	Autossômica Recessiva/ Autossômica Dominante	Anormalidades esqueléticas severas, esclera branca e possíveis defeitos neurológicos
<i>CREB3L1</i>	Substância induzida especificamente por astrócito antigo (OASIS; também conhecido como CR3L1)	XVI	Autossômica Recessiva	Deformidades ósseas severas
<i>SPARC</i>	SPARC (também conhecido como osteonectina)	XVII	Autossômica Recessiva	Fragilidade óssea severa progressiva
<i>MBTPS2</i>	Protease do sitio 2 do fator de transcrição ligado à membrana (S2P)	XVIII	Ligada ao X recessiva	Deformidade esquelética moderada a severa, esclera azul claro, escoliose e deformidades peitorais

Fonte: adaptada de Marini et al. (2017).

Esses dados evidenciam uma grande dificuldade de classificação dos pacientes diagnosticados, sendo que, nos casos em que não há diagnóstico genético, adota-se a classificação de Sillence et al. (1979).

Após a última década de avanços em Genética, duas abordagens para classificação da OI se desenvolveram: uma abordagem clínica na qual os novos tipos recessivos da doença são classificados dentre os tipos observados por Sillence et al. (1979), e uma abordagem genética-funcional, na qual os tipos I a IV, de Sillence et al. (1979) (osteogênese imperfeita clássica) são reservados para mutações nos genes *COL1A1* ou *COL1A2*, e as alterações em novos genes são classificados como novos tipos (MARINI et al., 2017).

### 2.2.3. Genes

Mais de 80% dos casos de OI são devidos a mutações nos genes *COL1A1* ou *COL1A2*, que codificam a  $\alpha 1$  e  $\alpha 2$  que são cadeias de colágeno tipo I, com padrão de herança autossômica dominante. Em geral, as mutações em genes não relacionados com o colágeno são associadas com formas recessivas de OI, com exceção no gene *IFITM5*, que codifica uma proteína transmembrana dos osteoblastos que podem regular os primeiros passos de mineralização. Outra exceção é observada com uma mutação no gene *MBTPS2*, associado à herança ligada ao X recessivo. Geralmente, mutações que causam perda de função em homozigotos ou heterozigotos resultam na diminuição severa ou na ausência de produção de proteína. Os genes com padrão de herança recessiva de OI podem ser geralmente categorizados com base nos caminhos celulares em que suas funções moleculares são executadas. Os genes recessivos de OI que são envolvidos na biossíntese do colágeno, modificação pós-traducional e processamento, incluem proteína associada a cartilagem (*CRTAP*), *P3H1*, *PPIB*, *FKBP10*, *SERPINH1*, *PLOD2*, *TMEM38B* e *BMP1*. As mutações em *SP7* e *WNT1* resultam em defeitos no desenvolvimento e atividade dos osteoblastos. A delimitação de mutações no fator de transcrição ligado à membrana, peptidase site 2 (*MBTPS2*) e *CREB3L1*, realça a importância crítica da proteólise intra-membrana regulada (*RIP*) no desenvolvimento esquelético. A compreensão melhorada da diferenciação osteoblástica e mineralização da matriz óssea é derivada de estudos de OI causados por mutações em *IFITM5* e *SERPINF1*. Geralmente, a identificação de

novos genes causadores de OI tem ampliado o escopo das vias celulares e biológicas para a patogênese da OI (KANG; ARYAL A.C.; MARINI, 2017).

### 2.3. NEXT GENERATION SEQUENCING

Com a expansão dos estudos de genômica, a demanda por tecnologias revolucionárias para a obtenção de informação genômica rápida, barata e precisa tornou-se proeminente. Esse desafio impulsionou o desenvolvimento dos equipamentos para sequenciamento de nova geração (MOROZOVA; MARRA, 2008).

A produção de dados em larga escala e a baixo custo, na ordem de milhões de leituras de sequências de DNA em uma única corrida, é a principal vantagem sobre os métodos convencionais. Desde que foi introduzido no mercado em 2005, as tecnologias de NGS têm tido enorme impacto na pesquisa genômica. Tecnologias de nova geração têm sido usadas para aplicações convencionais como sequenciamentos de genomas e para novas aplicações não exploradas anteriormente pelo sequenciamento Sanger, como a genômica funcional com foco na transcriptômica e epigenômica (SCHUSTER, 2008).

O sequenciamento de nova geração permite sequenciar o DNA em uma velocidade maior, permitindo, dessa forma, grandes avanços científicos e aplicações biológicas inéditas (MARDIS, 2008).

A variedade de características do NGS torna possível que múltiplas plataformas coexistam no mercado, com algumas tendo vantagens claras sobre outras, para aplicações específicas. Dentre as tecnologias de nova geração comercialmente disponíveis hoje destacam-se as plataformas 454, da Roche, HiSeq 2000, da Illumina e SOLiD, da *Applied Biosystems*. As tecnologias de sequenciamento evoluíram rapidamente e, no decorrer de 2011, novas plataformas foram lançadas; destaque para o Ion Torrent, da ThermoFisher, e o MiSeq, da Illumina (METZKER, 2010).

A plataforma Illumina utiliza uma metodologia similar ao método de Sanger, com nucleotídeos terminadores marcados com diferentes fluoróforos. Os fragmentos são ligados à *flowcell* (placa de vidro) e a emissão da fluorescência após cada nucleotídeo inserido é registrada para a obtenção da leitura (MOROZOVA; MARRA, 2008).

### **3. OBJETIVOS**

#### **3.1 OBJETIVO GERAL**

Identificar novos genes associados à Osteogênese Imperfeita por meio de recursos de bioinformática.

#### **3.2 OBJETIVOS ESPECÍFICOS**

- Pesquisar genes em bancos de dados genômicos envolvidos com as vias metabólicas de formação óssea, cartilaginosa e de colágeno;
- Selecionar uma lista de genes candidatos na plataforma customizada de NGS;
- Realizar busca de mutações em genes relacionados com a formação óssea e cartilaginosa em pacientes com Osteogênese Imperfeita que não possuam mutações identificadas em genes conhecidos;
- Verificar se as variações genéticas identificadas são potencialmente patogênicas.

## 4. METODOLOGIA

### 4.1. AMOSTRA

As amostras selecionadas para este trabalho foram provenientes de pacientes atendidos no Hospital Infantil Estadual Nossa Senhora da Glória, em Vitória/ES, centro de referência em tratamento da Osteogênese Imperfeita. Entre os pacientes submetidos ao sequenciamento de DNA por NGS, quatro deles apresentaram diagnóstico clínico para a OI, porém, resultados não conclusivos de genética molecular em NGS para os genes *P3H1*, *CRTAP*, *PPIB*, *FKBP10*, *WNT1*, *SP7* e *SERPINF1*, associados à OI. Portanto esses quatro pacientes foram selecionados para essa pesquisa (P.1, P.2, P.3 e P.4). A caracterização dos pacientes estudados quanto ao sexo, idade, tipo de herança genética, tipo de OI e dados clínicos estão descritos na tabela 3. Os heredogramas dos pacientes estão representados na Figura 3.

O quadro clínico desses pacientes foi avaliado segundo a classificação de Sillence et al. (1979), pois as alterações genéticas desses pacientes não eram conhecidas para que pudesse ser aplicada a classificação ampliada proposta por Marini et al. (2017).

Tabela 3: Caracterização dos pacientes estudados.

<b>Pacientes</b>	<b>P.1</b>	<b>P.2</b>	<b>P.3</b>	<b>P.4</b>
<b>Gênero</b>	Masculino	Masculino	Feminino	Feminino
<b>Idade (anos)</b>	17	15	11	9
<b>Herança</b>	Caso Isolado	Caso Isolado	Caso Isolado	Caso Familiar
<b>Classificação Sillence et al. (1979)</b>	Tipo I	Tipo III	Tipo IV	Tipo III
<b>Deformidades</b>	Membros superiores, leve a moderada	Membros superiores e inferiores, moderada a grave	Coluna, tórax, moderada em ossos longos não detalhados	Membro inferior, leve a moderada
<b>Esclera</b>	Normal	Azul ao nascimento	Normal	Cinza
<b>Deficiência intelectual</b>	Sim	Sim	Não	Sim
<b>Audição</b>	Normal	Normal	Normal	Normal
<b>Visão</b>	Normal	Normal	Normal	Normal
<b>Mobilidade</b>	Normal	Claudicação	Não anda	Não detalhada
<b>Informações adicionais</b>	Dentinogênese imperfeita	Dentinogênese imperfeita	Hipermobilidade das articulações, problemas cardíacos	Dentinogênese imperfeita

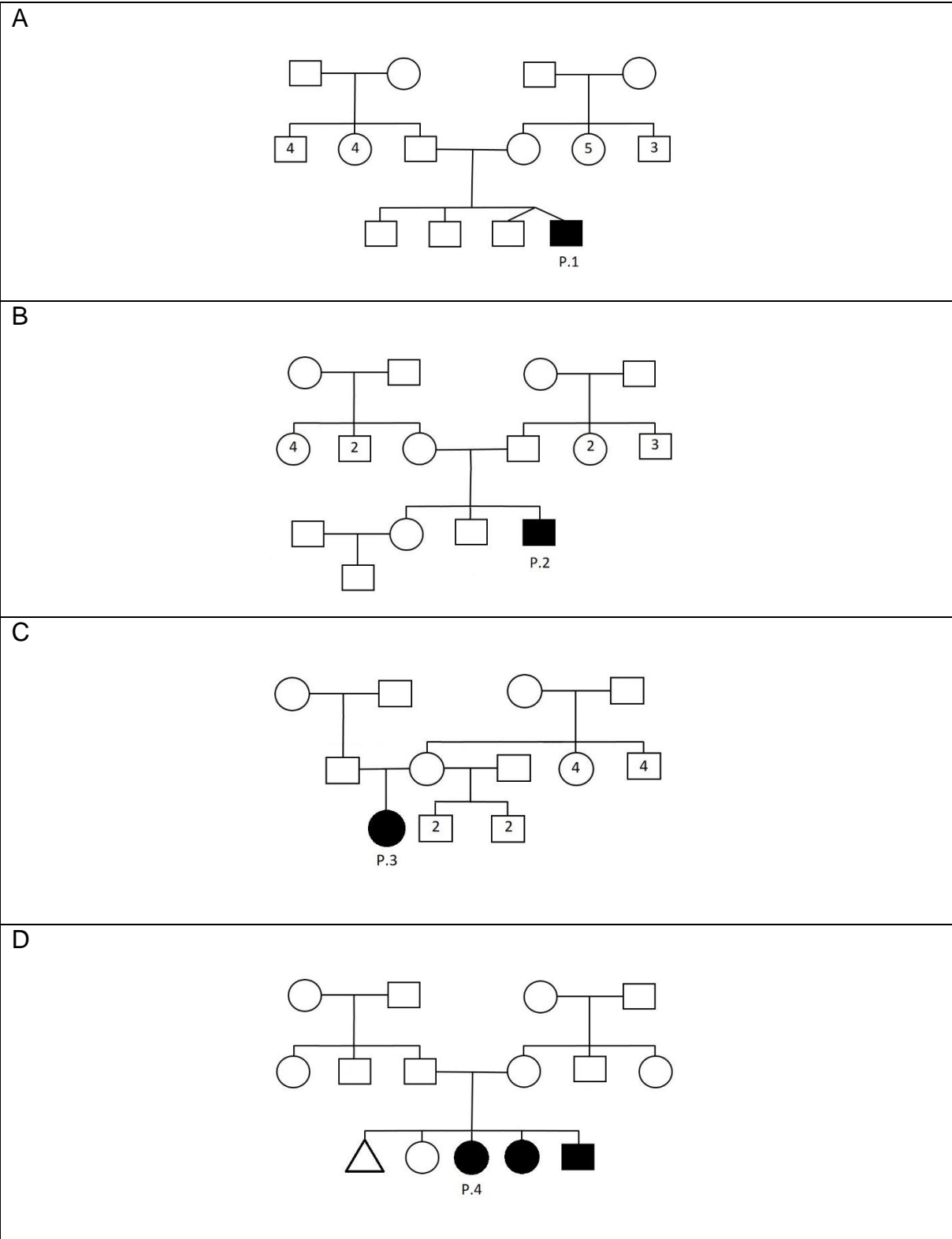


Figura 3: Heredogramas dos pacientes P.1 (A), P.2 (B), P.3 (C) e P.4 (D).



## 4.2. SELEÇÃO DOS PACIENTES ESTUDADOS

Após o tratamento dos dados provenientes do NGS, a análise dos genes conhecidos causadores de OI nos quatro pacientes estudados neste trabalho gerou resultados moleculares não conclusivos. Os quatro pacientes não apresentavam mutações nos genes *COL1A1* ou *COL1A2*. O paciente P.3 era portador de mutação no gene *FKBP10* em heterozigose. O paciente P.4 era portador de mutação em heterozigose no gene *P3H1*. Os genes *P3H1* e *FKBP10* estão relacionados com OI de herança autossômica recessiva. Assim, para causar a doença seriam necessárias duas mutações nestes genes. As mutações previamente identificadas nos genes *P3H1* e *FKBP10*, apesar de preditas como patogênicas, ainda não foram descritas na literatura e ainda não foram realizados estudos funcionais para validar se são realmente patogênicas.

## 4.3. MÉTODOS

Nesta pesquisa, foram selecionados quatro pacientes submetidos ao sequenciamento de 460 genes por NGS. Foram realizadas filtragens das variações genéticas nos genes selecionados de cada paciente, para identificar alterações raras, preditas como patogênicas e em regiões codificantes. Em seguida, no banco de dados genômicos *Ensembl*, foi feito um levantamento de genes candidatos, gerando 643 genes. A lista de genes candidatos foi comparada com os genes selecionados dos pacientes, restando 70 genes. Em seguida, foi realizada a exclusão de variações genéticas sinônimas e a análise de predição de mutações e proteínas. A apresentação dos métodos está esquematizada no fluxograma da figura 4.

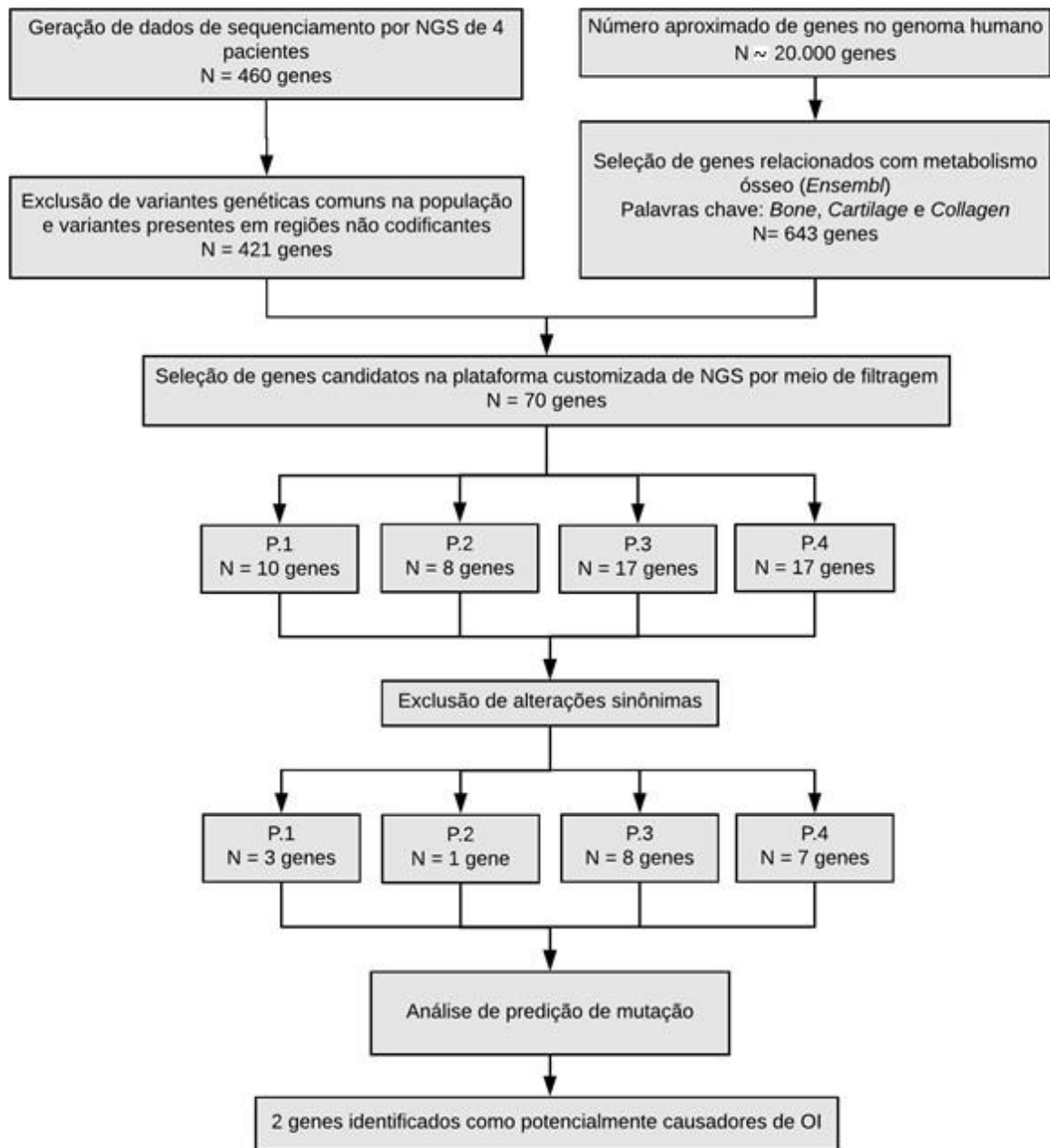


Figura 4: Fluxograma da metodologia.

## 4.4. ANÁLISE EXPERIMENTAL

### 4.4.1. Next Generation Sequencing

Foram obtidos dados genômicos de NGS de uma plataforma customizada que possuía 460 genes relacionados com doenças neuromusculares, doenças de genética do desenvolvimento e doenças do metabolismo ósseo. Também faziam parte desta plataforma os genes *COL1A1*, *COL1A2*, *FKBP10*, *CRTAP*, *P3H1*, *SP7*, e *SERPINH1*, conhecidos como causadores de OI.

Para a confecção do NGS foi gerada uma biblioteca a partir do DNA dos pacientes no painel Nextera – Illumina. Esta biblioteca foi validada verificando-se a quantificação da amplificação pelo Bioanalyzer com o *Kit High Sensivity DNA* e analisando o tamanho da biblioteca por Q-PCR com o kit KAPA Library Quantification no equipamento Real Time LighCycler da Roche, conforme recomendado pelo próprio fabricante. O sequenciamento ocorreu no equipamento MiSeq Illumina com 600 ciclos (2x300) e cobertura média de 120X.

Na análise de dados foi utilizado o programa BWA para realizar o alinhamento das sequências, encontrar os polimorfismos de nucleotídeos únicos (SNPs) e as sequências com inserção e/ou deleção (INDELs). Para o realinhamento das sequências e geração da lista de SNPs utilizou-se os programas Picard e o GATK.

Como controles foram utilizados dados genéticos de indivíduos fornecidos pelo grupo colaborador da USP. Esta etapa do trabalho foi feita em colaboração do Laboratório de Genética do Desenvolvimento e do Genoma Humano, da Universidade de São Paulo, com a técnica Meire Agüena, sob a supervisão das professoras Maria Rita Passos-Bueno e Mayana Zatz.

## 4.5. ANÁLISE DE BIOINFORMÁTICA

### 4.5.1. Parâmetros de bioinformática utilizados na seleção das variantes sequenciadas

Visto que a OI é uma doença genética rara, foi feito o processo de filtragem para excluir genes que possuísem variações genéticas comuns na população. Essas filtragens foram feitas utilizando-se os dados fornecidos pela plataforma do NGS, que estão descritos a seguir:

O *Allele count* (AC) indica a quantidade de alelos com alteração na população estudada. Quanto maior for o número de alelos contados, maior a chance de que a alteração seja um polimorfismo. Portanto, ao excluir os AC maiores que 10 excluimos os possíveis polimorfismos e mantivemos as possíveis mutações, visto que, para ser mutação, a alteração deve ser pouco frequente.

Outra informação do NGS que foi utilizada foi o AVSIFT. O SIFT usa várias informações de alinhamento para prever substituições toleradas e deletérias para cada posição da sequência de consulta. O SIFT é um procedimento de vários passos que (1) busca sequências semelhantes, (2) escolhe sequências intimamente relacionadas que podem compartilhar função semelhante à sequência de consulta, (3) obtém o alinhamento dessas sequências escolhidas e (4) calcula probabilidades normalizadas para todas as possíveis substituições do alinhamento. As posições com probabilidades normalizadas inferiores a 0,05 apresentam grande possibilidade de serem patogênicas, e aquelas cujos valores são maiores ou iguais a 0,05, prevê-se que sejam benignas (SIFT, 2018).

Portanto, como esta pesquisa tem como objetivo encontrar mutações patogênicas relacionadas à OI, foram selecionadas apenas mutações cujo AVSIFT fosse inferior a 0,05 e as que não possuísem informação (NA).

A coluna “Func.refGene” da planilha do NGS indica a localização do fragmento de DNA analisado. Como a pesquisa tem por objetivo encontrar mutações em genes ainda não associados à OI, decidiu-se buscar as mutações encontradas em éxons, que são as porções do gene que codificam a sequência de um polipeptídeo ou uma molécula funcional de RNA. Dessa forma, foram selecionadas apenas as alterações presentes em éxons.

“Gene.refGene” é a coluna que contém os nomes dos genes e, por fim, a coluna “ExonicFunc.refGene” classifica o tipo de alteração genética (sinônima, não-sinônima, *frameshift*, *non-frameshift*, *stopgain* ou *stoploss*).

#### 4.5.2. Etapas de filtragem dos dados

Para a seleção dos genes de interesse, foi realizada uma sequência de cinco etapas de filtragem, utilizando o programa Microsoft Office Excel 2013, na ordem apresentada a seguir:

A partir dos dados gerados pelo sequenciamento por NGS (N = 460 genes):

- 1: Foram excluídas as linhas cuja coluna “AC” fosse superior a 10;
- 2: Foram excluídas as linhas cuja coluna “AVSIFT” fosse superior a 0,05;
- 3: Foram selecionadas somente as linhas com “Func.refGene” iguais a “Exonic” (éxon) e excluídas todas as demais linhas;
- 4: Foram excluídos da análise todos os dados gerados por NGS não relacionados aos genes indicados na lista gerada pelo *Ensembl*. Desta forma obteve-se exclusivamente os dados de genes relacionados à formação óssea e cartilaginosa;
- 5: Foram excluídas as mutações sinônimas.

As etapas de filtragem de dados estão exemplificadas no anexo 1.

As etapas 1 a 3 foram realizadas mediante a aplicação de uma fórmula única, representada a seguir:

“=SE('Sequenciamento NGS'!\$AH<sub>n</sub>>10;"";SE(OU('Sequenciamento NGS'!\$Y<sub>n</sub><=0,05; 'Sequenciamento NGS'!\$Y<sub>n</sub>="NA");SE('Sequenciamento NGS'!\$F<sub>n</sub>="exonic"; 'Sequenciamento [NGS'!A<sub>n</sub>;"";"")))” onde:

n – número da linha (registro contendo os dados de cada alteração genética);

AH<sub>n</sub> - identifica a coluna onde estão os resultados do parâmetro “AC”;

Y<sub>n</sub> – identifica a coluna onde estão os resultados do parâmetro “AVSIFT”

$F_n$  – identifica a coluna onde está a informação acerca do local onde se encontra a cadeia de nucleotídeos alterada, indicada no NGS pelo nome “Func.refGene”

O resultado dessa filtragem foi a exclusão de todos os registros cujos parâmetros, analisados independentemente, apresentaram o valor “AC” superior a 10, o valor “AVSIFT” maior que 0,05 ou o campo “Func.refGene” com registro diferente de “exonic”.

#### 4.5.3. Seleção de genes candidatos

Após fazer a filtragem dos genes selecionando aqueles com maior probabilidade de serem patogênicos nas etapas 1 a 3, foi realizada a etapa 4.

Com o objetivo de selecionar genes candidatos que estivessem relacionados com a via do metabolismo ósseo no banco de dados genômico *Ensembl* (<http://www.ensembl.org/index.html>), foi realizada uma filtragem de todos os genes conhecidos utilizando as palavras chave: *bone*, *cartilage* e *collagen*. Além disso, foram aplicados filtros para “apenas pesquisando em humanos” e “apenas pesquisando genes”. Essa busca gerou uma lista de 643 genes.

Paralelamente, foram obtidos resultados de sequenciamento de 460 genes dos pacientes por NGS.

Após esta etapa, foi feita uma filtragem no Microsoft Office Excel 2013 com o intuito de manter os genes em comum. A fórmula utilizada, transcrita a seguir, replicava, em uma nova planilha, o conteúdo das linhas da planilha original cujo campo “Gene” apresentasse conteúdo igual ao de alguma das células da planilha obtida a partir da pesquisa no *Ensembl*:

“=SE(CONT.SE('Genes pesquisados'!\$A:\$A;'Exclui linhas vazias limpa gene'!\$I<sub>n</sub>)=0;"";'Exclui linhas vazias limpa gene'!A<sub>n</sub>)”, onde:

$n$  – número da linha (registro contendo os dados de cada alteração genética);

$A$  – coluna da planilha “Genes pesquisados” onde estavam relacionados os genes obtidos a partir da pesquisa no *Ensembl*;

In – célula de uma determinada linha (registro) onde estava informado o gene no qual a alteração sob análise foi identificada;

O resultado dessa filtragem foi uma planilha com todos os registros que continham, em seu campo “Gene”, algum dos 643 genes da lista obtida no Ensembl, contendo as alterações identificadas em 70 genes, todos eles coincidentes com os genes constantes da listagem obtida a partir da pesquisa no *Ensembl*. Os outros registros, descartados para essa pesquisa, não estavam contidos na listagem de genes obtida a partir da pesquisa no Ensembl.

Dos 460 genes analisados por NGS, 70 (10,9%) estavam entre os genes identificados no *Ensembl*, relacionados com a via do metabolismo ósseo.

#### 4.5.4. Exclusão das mutações sinônimas

Na etapa 5, foram excluídas as mutações sinônimas individualmente, sem a utilização de fórmulas.

Com essa filtragem, foram obtidas informações de alterações genéticas relacionadas a osso, cartilagem e colágeno; potencialmente raras, potencialmente patogênicas, em região de éxon e não-sinônimas.

Os resultados foram agrupados por paciente para que cada caso pudesse ser analisado individualmente.

#### 4.5.5. Análise de predição de mutações e proteínas

Cada alteração encontrada foi avaliada nos *softwares* de predição:

- *Mutation taster* (<http://www.mutationtaster.org/>),
- *PolyPhen-2* (<http://genetics.bwh.harvard.edu/pph2/index.shtml>) e
- *PROVEAN* ([http://provean.jcvi.org/seq\\_submit.php](http://provean.jcvi.org/seq_submit.php)).

Além disso, foram realizadas buscas no banco de dados do ClinVar (<https://www.ncbi.nlm.nih.gov/clinvar/>) para verificar se a alteração já estava descrita na literatura.

O *MutationTaster* emprega um classificador para eventualmente prever o potencial de uma alteração ser causadora de doença. O classificador é alimentado com o resultado de todos os testes e as características das alterações e calcula as probabilidades de que a alteração seja uma mutação causadora de doença ou um polimorfismo inofensivo. Para esta previsão, as frequências de todos os recursos únicos para mutações / polimorfismos de doenças conhecidas foram estudadas em um grande conjunto de treinamento composto por mais de 390.000 mutações de doenças conhecidas no *Human Gene Mutation Database* (HGMD) e mais de 6.800.000 SNPs inofensivos e polimorfismos de Indel do Projeto 1.000 Genomas. (<http://www.mutationtaster.org/info/documentation.html>).

O Projeto 1.000 Genomas fornece uma descrição abrangente de humanos com variações genéticas comuns aplicando o sequenciamento de todo o genoma em um conjunto diverso de indivíduos de múltiplas populações. (AUTON et al., 2015).

O *Exome Aggregation Consortium* (ExAC) é uma agregação em larga escala de dados genômicos humanos, ou seja, uma associação de pesquisadores que procuram agregar e harmonizar dados de sequenciamento do exoma de uma variedade de projetos de sequenciamento em larga escala e disponibilizar dados resumidos para a comunidade científica em geral. O conjunto de dados fornecido neste site abrange 60.706 indivíduos não relacionados, sequenciados como parte de vários estudos específicos de doenças e de genética populacional. Portanto, este conjunto de dados deve servir como um conjunto de referência útil de frequências de alelos para estudos de doenças graves (<http://exac.broadinstitute.org/about>).

Na base de dados do *Mutation taster* contém todos os polimorfismos de nucleotídeos únicos (SNPs) do banco de dados NCBI SNP (dbSNP). Se uma alteração foi encontrada mais de quatro vezes em homozigose nos 1.000 genomas, ela é automaticamente considerada como polimorfismo. O *Mutation Taster* também exhibe variantes de doenças conhecidas do dbSNP ClinVar. Se uma variante é marcada como provavelmente-patogênica ou patogênica no ClinVar, prevê-se que ela seja causadora de doença, ou seja, a alteração é classificada automaticamente como causadora de doença.



O *PolyPhen-2* (*Polymorphism Phenotyping v2*) é uma ferramenta que prevê o possível impacto de uma substituição de aminoácidos na estrutura e função de uma proteína humana usando considerações físicas e comparativas diretas. Alguns dos destaques são: *Pipeline* de alinhamento de sequências múltiplas de alta qualidade, classificador probabilístico baseado no método de aprendizagem mecânica e foi otimizado para análises de alto rendimento dos dados de sequenciamento da próxima geração (<http://genetics.bwh.harvard.edu/pph2/index.shtml>).

O *PROVEAN* apresenta um índice de alinhamento versátil como uma métrica para prever os efeitos nocivos de variações não limitadas a substituições de aminoácidos simples, mas também inserções no quadro, deleções e múltiplas substituições de aminoácidos. Esta pontuação baseada em alinhamento mede a mudança na similaridade de sequência de uma sequência de consulta para um homólogo de sequência de proteína antes e depois da introdução de uma variação de aminoácido na sequência de consulta. A ferramenta PROVEAN está disponível on-line em [http://provean.jcvi.org/seq\\_submit.php](http://provean.jcvi.org/seq_submit.php).

Todas as alterações também foram pesquisadas no banco de dados do NCBI para verificar se já haviam sido publicadas e se já tinham sido feitos estudos clínicos verificando a patogenicidade de cada alteração.

Para a predição da estrutura secundária da proteína do principal gene encontrado e do gene com mutação, foi utilizado o site do Bloomsbury Centro de Bioinformática (<http://bioinf.cs.ucl.ac.uk/psipred/>).

## 5. RESULTADOS

A metodologia adotada permitiu a identificação de variantes genéticas nos quatro pacientes, como descritos nas tabelas 4, 5, 6 e 7. Os resultados da análise destes genes permitiram a identificação de dois genes potencialmente causadores de OI: *ALX1* e *RYR1*. A análise de predição de proteína *ALX1 Homeobox 1* da sequência selvagem e sequência mutada demonstrou modificações da estrutura secundária desta proteína, como observado nas figuras 9 e 10. Essa mesma análise não foi possível para a proteína codificada do gene *RYR1*, pois o sistema de predição de proteínas não foi capaz de analisar esta proteína devido ao seu extenso tamanho.

Tabela 4: Alterações genéticas encontradas no paciente 1.

Paciente: P.1									
Gene	Alelos	Alteração no DNA	Tipo	AVSift	Alteração na proteína	Predição de mutação no <i>Mutation Taster</i>	Predição de mutação de SNPs no <i>PolyPhen-2</i>	Predição de proteína PROVEAN	Registro da mutação no ClinVar-NCBI
<i>ALX1</i>	0/1	NM_006982: exon3: c.610_630del	Deleção <i>Non-frameshift</i>	NA	p.204_210del	Causador de doença – Longa deleção	Não se aplica	Deletério	Não consta
<i>COL6A1</i>	0/1	NM_001848: exon19: c. <b>G</b> 1316 <b>A</b>	Não-sinônima SNV	00,05	p.R439Q	Causador de doença	Provavelmente danoso	Neutro	Benigno/ Provavelmente benigno
<i>COL6A2</i>	0/1	NM_001849: exon28: c. <b>A</b> 3043 <b>C</b>	Não-sinônima SNV	NA	p.I1015L	Polimorfismo	Possivelmente danoso	Neutro	Benigno/ Provavelmente benigno

NA = não avaliado; 0/1=heterozigoto; SNV = Variação de nucleotídeo único

Tabela 5 - Alterações genéticas encontradas no paciente 2.

Paciente: P.2									
Gene	Alelos	Alteração no DNA	Tipo	AVSift	Alteração na proteína	Predição de mutação no <i>Mutation Taster</i>	Predição de mutação de SNPs no <i>PolyPhen-2</i>	Predição de proteína PROVEAN	Registro da mutação no ClinVar-NCBI
COL6A3	0/1	NM_057166: exon38: c. <b>C7156T</b>	Não-sinônima SNV	NA	p.R2386C	Polimorfismo	Benigno	Neutro	Não consta
		NM_057167: exon40: c. <b>C8359T</b>	Não-sinônima SNV	NA	p.R2787C	Polimorfismo	Benigno	Neutro	Não consta
		NM_004369: exon41: c. <b>C8977T</b>	Não-sinônima SNV	NA	p.R2993C	Polimorfismo	Possivelmente danoso	Neutro	Não consta

NA = não avaliado; 0/1=heterozigoto; SNV = Variação de nucleotídeo único

Tabela 6 - Alterações genéticas encontradas no paciente 3.

Paciente: P.3									
Gene	Alelos	Alteração no DNA	Tipo	AVSift	Alteração na proteína	Predição de mutação no <i>Mutation Taster</i>	Predição de mutação de SNPs no <i>PolyPhen-2</i>	Predição de proteína PROVEAN	Registro da mutação no ClinVar-NCBI
<i>ALPL</i>	0/1	NM_001177520: exon10: c.1343delG	Deleção <i>Non-frameshift</i>	NA	p.X448X	Causador de doença	Não se aplica	Não se aplica	Não consta
<i>AXIN2</i>	0/1	NM_004655: exon8: c.2013_2024del	Deleção <i>Non-frameshift</i>	NA	p.670_673del	Causador de doença	Não se aplica	Deletério	Benigno
		NM_004655: exon6: c.G1615A	Não-sinônima SNV	NA	p.V539M	Causador de doença	Provavelmente danoso	Neutro	Benigno/ Provavelmente benigno
		NM_004655: exon8: c.C1952T	Não-sinônima SNV	NA	p.S651L	Polimorfismo	Benigno	Neutro	Provavelmente benigno
<i>COL6A3</i>	0/1	NM_004369: exon10: c.G4727A	Não-sinônima SNV	00,02	p.R1576Q	Polimorfismo	Benigno	Neutro	Benigno/ Provavelmente benigno
<i>COMP</i>	0/1	NM_000095: exon5: c.C396G	Não-sinônima SNV	NA	p.N132K	Polimorfismo	Benigno	Neutro	Não consta
<i>FKBP10</i>	0/1	NM_021939: exon1: c.A179C	Não-sinônima SNV	00,01	p.Q60P	Causador de doença	Provavelmente danoso	Deletério	Não consta
<i>GDF5</i>	0/1	NM_000557: exon1: c.C497A	Não-sinônima SNV	00,04	p.P166H	Polimorfismo	Possivelmente danoso	Neutro	Benigno

GNAS	0/1	NM_001077490: exon1: c.992_993ins <b>GCA GCCCCCT</b>	Inserção Non- frameshift	NA	G331_Q332in sGQPL	Polimorfismo	Não se aplica	Neutro	Não consta
		NM_080425: exon1: c.1179_1180ins <b>GC AGCCCCCT</b>	Inserção Non- frameshift	NA	R393_A394ins RAAP	Polimorfismo	Não se aplica	Neutro	Não consta
USH2A	0/1	NM_206933: exon71: c. <b>T15377C</b>	Não- sinônima SNV	NA	p.I5126T	Polimorfismo	Benigno	Neutro	Benigno

NA = não avaliado; 0/1=heterozigoto; SNV = Variação de nucleotídeo único

Tabela 7 - Alterações genéticas encontradas no paciente 4

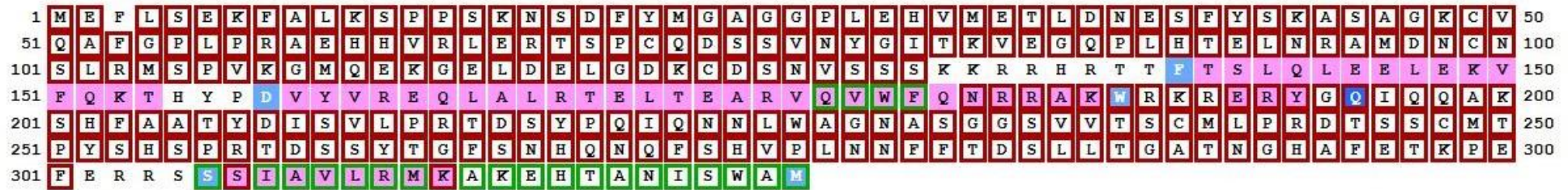
Paciente: P.4									
Gene	Alelos	Alteração no DNA	Tipo	AVSift	Alteração na proteína	Predição de mutação no <i>Mutation Taster</i>	Predição de mutação de SNPs no <i>PolyPhen-2</i>	Predição de proteína PROVEAN	Registro da mutação no ClinVar-NCBI
AXIN2	0/1	NM_004655: exon8: c. <b>2013_2024del</b>	Deleção Non- frameshift	NA	p.670_673del	Causador de doença	Não se aplica	Deletério	Benigno
		NM_004655: exon6: c. <b>C1573G</b>	Não- sinônima SNV	NA	p.P525A	Causador de doença	Benigno	Neutro	Benigno
COL5A2	0/1	NM_000393: exon22: c. <b>C1454A</b>	Não- sinônima SNV	00,00	p.P485Q	Causador de doença	Provavelmente danoso	Deletério	Benigno/ Provavelmente benigno

<i>EVC</i>	0/1	NM_153717: exon3: c. <b>C341T</b>	Não- sinônima SNV	00,00	p.A114V	Polimorfismo	Provavelmente danoso	Deletério	Benigno
<i>P3H1</i>	0/1	NM_001146289: exon14: c. <b>G2024T</b>	Não- sinônima SNV	00,00	p.W675L	Causador de doença	Provavelmente danoso	Deletério	Não consta
<i>RYS1</i>	0/1	NM_001042723: exon44: c. <b>C7102G</b>	Não- sinônima SNV	00,00	p.L2368V	Causador de doença	Provavelmente danoso	Deletério	Não consta
<i>TRIP11</i>	0/1	NM_004239: exon11: c. <b>G2134A</b>	Não- sinônima SNV	00,01	p.E712K	Causador de doença	Possivelmente danoso	Neutro	Benigno/ Provavelmente benigno
<i>USH2A</i>	0/1	NM_206933: exon48: c. <b>G9430A</b>	Não- sinônima SNV	NA	p.D3144N	Polimorfismo	Benigno	Neutro	Benigno
		NM_206933: exon47: c. <b>A9296G</b>	Não- sinônima SNV	NA	p.N3099S	Polimorfismo	Benigno	Neutro	Benigno

NA = não avaliado; 0/1=heterozigoto; SNV = Variação de nucleotídeo único

### Mapa de Estrutura Secundária

Predições apresentadas são codificadas em cores dentro da sequência conforme a legenda apresentada abaixo.



Legenda:

Hélice

Folha-Beta

Desordenado

Ligação de proteína desordenada

Domínio DomPred

Domínio DomSSEA

Notações:

M

L

E

E

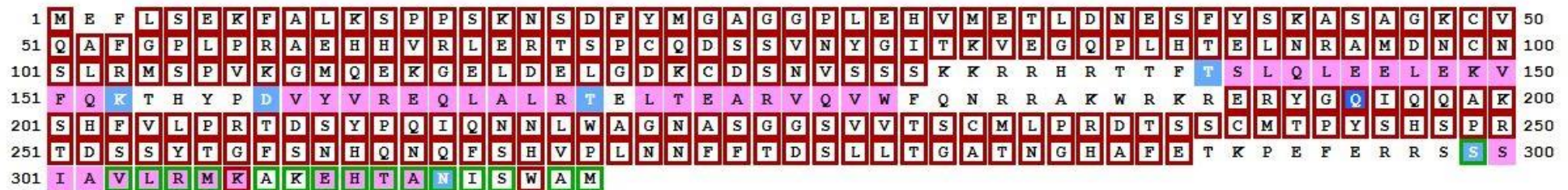
A

D

Figura 5: ALX1 sequência selvagem.

### Mapa de Estrutura Secundária

Predições apresentadas são codificadas em cores dentro da sequência conforme a legenda apresentada abaixo.



Legenda:

Hélice

Folha-Beta

Desordenado

Ligação de proteína desordenada

Domínio DomPred

Domínio DomSSEA

Notações:

M

L

E

E

A

D

Figura 6: ALX1 Sequência mutada.

## 6. DISCUSSÃO

### 6.1. VARIAÇÕES GENÉTICAS IDENTIFICADAS

#### 6.1.1. Paciente P.1

Neste paciente, que apresenta sintomas brandos de OI, foram encontradas mutações em heterozigose nos genes: *ALX1*, *COL6A1* e *COL6A2*.

##### 6.1.1.1. Variante no gene *ALX1*

Foi identificada uma deleção *non-frameshift* de 21 nucleotídeos, resultando na perda de sete aminoácidos, no gene *ALX1* (*ALX homeobox 1*) que codifica uma proteína associada à cartilagem (ZHAO et al., 1993). A mutação identificada, ainda não descrita na literatura, foi predita como causadora de doença pelo *Mutation Taster* e pelo *PROVEAN*. O segmento deletado é idêntico ao de várias espécies de vertebrados, como: *Ptrogodytes* (chimpanzé), *Mmulatta* (macaco Rhesus), *Mmusculus* (rato doméstico), *Ggallus* (galinha), *Drerio* (peixe-zebra) e *Xtropicalis* (rã), sugerindo que a sequência deletada desta proteína possui importância relevante em espécies relacionadas evolutivamente (SCHWARZ et al., 2014). Além disso, essa variação genética não foi encontrada no ExAC e nem nos 1.000 genomas, sugerindo se tratar de uma alteração rara, compatível com a ocorrência de doenças genéticas de baixa prevalência na população. Não foram descritos na literatura casos de mutações heterozigóticas no gene *ALX1*. Porém já foi relatado no gene *ALX4*, um gene da mesma família, que mutações em heterozigose geraram sintomas leves da doença genética displasia frontonasal (FND) (BERTOLA et al., 2013).

##### 6.1.1.2 Variantes no gene *COL6A1*

Foi encontrada uma mutação em heterozigose, predita como causadora de doença e provavelmente danosa no gene *COL6A1*, que participa da formação do colágeno VI (LAMANDÉ; BATEMAN, 2018). Esta alteração genética resultou em uma substituição



do aminoácido arginina para glutamina, o que poderia provocar mudanças nas características da proteína codificada.

Entretanto, esta mesma mutação já foi relatada em heterozigose em 27 indivíduos no banco de dados dos 1.000 genomas e em 189 casos relatados no ExAC. Além disso, esta mutação teve predição de proteína neutra pelo PROVEAN e foi divulgada com provável significância clínica benigna no banco de dados do NCBI pelo EGL (*Emory Molecular Genetics Laboratory*) em 2012 e pelo *Illumina Clinical Services Laboratory* (ICSL) em 2016 - ([https://www.ncbi.nlm.nih.gov/clinvar/variation/93810/#cvrec\\_primary93810](https://www.ncbi.nlm.nih.gov/clinvar/variation/93810/#cvrec_primary93810)).

Assim, apesar dos programas de predição sugerirem que esta seja uma alteração patogênica, as evidências biológicas apoiam a hipótese de que esta mutação não seja responsável por causar doença no ser humano.

#### 6.1.1.3. Variante no gene COL6A2

Foi identificada uma mutação em heterozigose, predita como polimórfica pelo *Mutation taster* e como possivelmente danosa pelo *PolyPhen-2*, no gene COL6A2, que participa da formação do colágeno VI (LAMANDÉ; BATEMAN, 2018). Houve uma alteração do aminoácido isoleucina para leucina, ambos de caráter neutro. Esta mutação possui 96 casos relatados em heterozigose no banco de dados dos 1.000 genomas e 685 casos relatados no ExAC. Esta alteração já foi descrita no banco de dados do NCBI sete vezes, entre 2012 e 2016, e apresentou significância clínica benigna em vários bancos de dados genômicos [EGL em 2012, *American College of Medical Genetics and Genomics* (ACMG) em 2015, ICSL em 2016 e *General Variant Classification Assertion Criteria* (GeneDx) em 2016] - (<https://www.ncbi.nlm.nih.gov/clinvar/variation/93952/>).

#### 6.1.1.4 Conclusão sobre as análises das variantes do Paciente P.1

Como o paciente P.1 não apresenta mutações conhecidas, a mutação no gene *ALX1* pode ser a causa das manifestações clínicas desse paciente. Com base nos resultados obtidos e nas análises da literatura, este paciente é portador de uma

mutação no gene *ALX1*, potencialmente patogênica, enquanto as alterações nos genes *COL6A1* e *COL6A2* são, provavelmente, benignas.

#### 6.1.2. Paciente P.2

O paciente P.2 apresentou mutações em heterozigose no gene *COL6A3*, que codifica uma proteína que participa da formação do colágeno VI (LAMANDÉ; BATEMAN, 2018). Foram identificadas mutações de troca de nucleotídeo que resultaram na troca do aminoácido arginina (básico) para cisteína (neutro). Apesar dessas alterações não terem sido descritas no 1.000 genomas nem no NCBI e terem sido relatadas duas vezes no ExAC, elas foram preditas como polimórficas (*Mutation Taster*), benignas (*PolyPhen-2*) e neutras (*PROVEAN*). Apenas a variante NM\_004369 apresentou predição de possivelmente danosa no *PolyPhen-2*, porém a respectiva proteína foi predita como neutra. Assim, as estratégias *in silico* adotadas nesta pesquisa não identificaram mutações potencialmente patogênicas no paciente P.2.

#### 6.1.3. Paciente P.3

Neste paciente foram encontradas mutações em heterozigose nos genes: *ALPL*, *AXIN2*, *COL6A3*, *COMP*, *FKBP10*, *GDF5*, *GNAS* E *USH2A*.

Apesar de este paciente ser portador de uma mutação em um gene causador de OI previamente conhecido, o gene *FKBP10*, ele foi selecionado para a pesquisa porque este gene está relacionado com OI de herança autossômica recessiva e o paciente é portador de mutação em heterozigose neste gene, como confirmado neste trabalho.

Também foram encontradas, neste paciente, mutações nos genes *COL6A3*, *COMP*, *GDF5*, *GNAS* e *USH2A* preditas como polimórficas pelo *Mutation Taster* e neutras pelo *PROVEAN*. Além disso, as variantes dos genes *COL6A3*, *GDF5* e *USH2A* foram descritas pelo NCBI como de significâncias clínicas benignas. Por isso, esses genes, provavelmente, não estão contribuindo para as manifestações clínicas deste paciente.

#### 6.1.3.1. Variantes no gene *COL6A3*

O paciente apresentou uma mutação em heterozigose no gene *COL6A3*, que codifica uma proteína que participa da formação do colágeno VI (LAMANDÉ; BATEMAN, 2018). Essa mutação foi detectada pelo NGS em três diferentes referências. Nas referências NM\_057166 e NM\_057167, a mutação foi predita como provavelmente danosa pelo *PolyPhen-2*, polimórfica pelo *Mutation taster*, neutra pelo *PROVEAN* e não consta no banco de dados do NCBI. No caso do NM\_004369, foi predita como polimorfismo, benigno, neutro e foi relatado no NCBI como benigno. Assim, não foram identificadas mutações potencialmente patogênicas para o gene *COL6A3* no paciente P.3.

#### 6.1.3.2. Variantes no gene *AXIN2*

Mutações no gene *AXIN2*, em humanos, estão relacionadas a alguns tipos de câncer e a problemas de mineralização óssea (EHLEN et al., 2013; LIU et al., 2000). O paciente P.3 apresentou três mutações em heterozigose no gene *AXIN2*.

A primeira mutação consistiu em uma deleção *non-frameshift* de 12 nucleotídeos. Esta deleção foi previamente descrita em heterozigose em 101 casos nos 1.000 genomas e 737 no ExAC. Apesar de ela ser predita como causadora de doença por dois preditores (*Mutation taster* e *PROVEAN*), foi relatada com significância clínica benigna de acordo com GeneDx, 2014; ACMG, 2015 e Invitae Variant Classification (IVC), 2017.

A segunda mutação foi de substituição de um nucleotídeo, resultando em troca do aminoácido valina (neutro) para metionina (neutro). A predição da mutação foi de causadora de doença (*Mutation taster*) e provavelmente danoso (*PolyPhen-2*), porém foi relatada em heterozigose 67 vezes nos 1.000 genomas e 463 no ExAC e foi predita como neutra pelo *PROVEAN*. No *ClinVar* foi predita como benigna ou provavelmente benigna de acordo com GeneDx, 2013; ACMG, 2015; IVC, 2017 e ICSL, 2016.

A terceira mutação, que foi uma troca simples de base C>T, com mudança de aminoácido serina para leucina foi predita como polimorfismo. Sua significância clínica foi relatada no NCBI como provavelmente benigna por IVC, 2016 e GeneDx, 2017.

Portanto, apesar de duas das mutações em *AXIN2* terem sido preditas como potenciais causadoras de doença, estudos funcionais demonstraram sua significância clínica benigna. Por isso, essas alterações, provavelmente, não estão contribuindo para os sintomas deste paciente.

#### 6.1.3.3. Variantes no gene *ALPL*

Foi identificada uma mutação em heterozigose no gene *ALPL*, que codifica uma proteína que participa da formação dos ossos (WHYTE, 2017). A mutação foi de deleção de um nucleotídeo de guanina, predita como patogênica, sem alteração do quadro de leitura da proteína. Esta mutação ocorreu no códon de parada e alterou a região de *splicing*. Apesar de estas mutações não constarem no banco de dados do NCBI, nem nos 1.000 genomas, elas já foram relatadas em 34 casos em heterozigose no ExAC. Esta alteração foi avaliada usando-se a referência NM\_001177520, mas as análises de NGS também sugerem as referências NM\_000478 e NM\_001127501 para a mesma mutação. As análises de predição geraram os mesmos resultados nas três referências. Aproximadamente 70% de variações no gene *ALPL* identificadas no NCBI (ClinVar) estão associados à hipofosfatasia. Esta é uma doença genética rara que afeta, especialmente, crianças e provoca deformações e fraturas em algumas regiões do corpo e perda prematura dos dentes de leite, por defeito na mineralização óssea. A transmissão de formas graves é autossômica recessiva, enquanto formas mais leves podem ser transmitidas de forma autossômica dominante ou recessiva. A doença decorre de mutações no gene da fosfatase alcalina tipo hepática (*ALPL*), também chamado de gene *TNSALP*. (BACON; CROWLEY, 2018; WHYTE, 2016).

Portanto, como vários sintomas da hipofosfatasia são semelhantes aos sintomas da OI, a alteração genética deste paciente indica que ele pode ser portador de hipofosfatasia. Assim, são, no futuro, necessários estudos complementares para reavaliar o diagnóstico clínico.

#### 6.1.3.4. Conclusão sobre as análises das variantes do Paciente P.3

As análises deste trabalho sugerem que as mutações identificadas no paciente P.3, nos genes *AXIN2*, *COL6A3*, *COMP*, *GDF5*, *GNAS* e *USH2A*, são benignas.

Este paciente também é portador de mutações em heterozigose em dois genes: 1) no gene *FKBP10*, relacionado com OI com padrão de herança autossômica recessiva e 2) no gene *ALPL*, relacionado com hipofosfatasia, doença com sintomas de sobreposição com OI.

Assim, são necessários estudos complementares para esclarecer o diagnóstico clínico deste paciente, para que, em conjunto com os dados de genética molecular, seja possível contribuir com a ampliação do nosso conhecimento sobre as bases moleculares da doença investigada.

#### 6.1.4. Paciente P.4

No paciente P.4 foram encontradas mutações em heterozigose nos genes: *AXIN2*, *COL5A2*, *EVC*, *P3H1*, *RYR1*, *TRIP11* e *USH2A*.

Mutações no gene *P3H1* são a causa de OI de herança autossômica recessiva. Este trabalho confirmou a presença de apenas uma mutação em heterozigose neste paciente, gerando um resultado molecular inconclusivo. Esta alteração foi avaliada usando-se a referência NM\_001146289, mas as análises de NGS também sugerem as referências NM\_001243246 e NM\_022356 para a mesma mutação. As análises de predição nas diferentes referências geraram os mesmos resultados.

##### 6.1.4.1. Variantes no gene *AXIN2*

Foram encontradas duas mutações preditas como causadoras de doença no gene *AXIN2*. Uma é idêntica à do paciente P.3: deleção de 12 nucleotídeos *non-frameshift*, predita como causadora de doença, mas com significância clínica benigna de acordo com GeneDx, 2014; ACMG, 2015 e IVC, 2017.

A outra mutação encontrada neste gene foi de uma substituição do nucleotídeo C para G, que resultou na troca do aminoácido prolina para alanina. Esta mutação já foi citada e possui significância clínica benigna (GeneDx em 2013 e IVC em 2017).

#### 6.1.4.2. Variantes no gene *COL5A2*

Ocorreu uma mutação em heterozigose de substituição do nucleotídeo C para A, resultando em uma troca do aminoácido prolina para glutamina. Essa substituição apresentou predição de causadora de doença, provavelmente danosa e deletéria. Ela foi relatada em 6 indivíduos em heterozigose nos 1.000 genomas e 65 casos no ExAC. No entanto, a mutação consta no NCBI e possui significância clínica benigna (IVC, 2017) e provavelmente benigna (GeneDx, 2017; *Ambry Autosomal Dominant and X-linked criteria*, 2015 e ICSL, 2016).

#### 6.1.4.3. Variantes no gene *TRIP11*

Ocorreu uma mutação no gene *TRIP11* predita como causadora de doença e possivelmente danosa, porém ela foi relatada como neutra no *PROVEAN* e, no NCBI, como benigna e provavelmente benigna, sendo afastada a sua significância para os sintomas da OI. Apesar destas mutações não constarem no banco de dados do NCBI nem dos 1.000 genomas, elas já foram relatadas em 34 indivíduos no ExAC.

#### 6.1.4.4. Variante no gene *RYS1*

No presente estudo, ocorreu uma alteração de substituição do nucleotídeo C>G na posição 7102, resultando na substituição do aminoácido leucina por valina (p.L2368V). As predições de mutação desta alteração sugerem que ela seja patogênica. Esta alteração foi avaliada usando-se a referência NM\_001042723, mas as análises de NGS também sugerem a referência NM\_000540 para a mesma mutação. A análise de predição para a referência NM\_000540 foi relatada e está descrita como uma variante de significância incerta (ClinVar). Além disso foi predita como patogênica pelo *Mutation Taster* e pelo *PROVEAN*, enquanto que, pelo PolyPhen-2, foi predita como

benigna. Já para a variante NM\_001042723, a alteração foi considerada como patogênica por todos os preditores utilizados e não consta no banco de dados do ClinVar (NCBI). Apesar de ter sido encontrado um resultado diferente entre as variantes, a alteração é a mesma (p.L2368V), sugerindo que o paciente possua apenas uma mutação que foi avaliada pelo NGS por duas referências distintas. Por isso, essas análises sugerem que a variação encontrada no gene *RYR1* do paciente P.4 seja patogênica.

#### 6.1.4.5. Conclusão sobre as análises das variantes do Paciente P.4

De acordo com as análises de predição, as alterações encontradas nos genes *AXIN2*, *COL5A2*, *EVC*, *TRIP11* e *USH2A* são benignas.

Este paciente também é portador em heterozigose de mutações em dois genes: 1) no gene *P3H1* relacionado com OI com padrão de herança autossômico recessivo e 2) no gene *RYR1* relacionado com doenças musculares e relacionado também ao metabolismo ósseo.

## 6.2. GENES POTENCIALMENTE PATOGÊNICOS

### 6.2.1. Gene *ALX1*

O gene *ALX1*, também chamado de *CART1*, está relacionado à doença displasia frontonasal, que descreve uma série de anormalidades. Principais achados incluem verdadeiro hipertelorismo ocular; alargamento da raiz nasal; fenda facial média que afeta o nariz e / ou lábio superior e palato; falta de formação da ponta nasal e linha frontal em forma de V (UZ et al., 2010). A perda completa da função da proteína *ALX1* provoca uma ruptura grave do desenvolvimento craniofacial precoce. Ao contrário da perda de seu ortólogo murino, a perda de *ALX1* humana não resulta em defeitos do tubo neural; no entanto, afeta severamente a fusão orquestrada entre processos frontonais, nasomediais, nasolaterais e maxilares durante embriogênese em estágio inicial (UZ et al., 2010).

Recentemente, os pacientes com displasia frontonasal autossômica recessiva foram ligados a mutações de perda de função homozigóticas na família de genes *homeobox ALX*, que incluem genes *ALX1*, *ALX3* e *ALX4* (KAYSERILI et al., 2009; TWIGG et al., 2009; UZ et al., 2010). Esta família de genes pertence às homeoproteínas de classe emparelhada e desempenha um papel importante nos processos complexos de desenvolvimento craniano, bem como no fechamento do tubo neural e no desenvolvimento de membros (MCGONNELL et al., 2011). Em contraste, as mutações heterozigóticas de perda de função em *ALX4* foram originalmente associadas a defeitos simétricos de ossificação parietal isolados na interseção das suturas sagital e lambdoide, conhecidas como forames parietais ampliados, sem características faciais da displasia frontonasal (BERTOLA et al., 2013).

A proteína *ALX homeobox protein 1* desempenha a função de fator de transcrição de ligação ao DNA específico da sequência que se liga a sequências palindrômicas nos promotores e pode ativar ou reprimir a transcrição de um subconjunto de genes (CAI, 1998; GORDON et al., 1996). Muito provavelmente regula a expressão de genes envolvidos no desenvolvimento de estruturas craniofaciais derivadas do mesênquima. No início do desenvolvimento, ele desempenha um papel na sobrevivência do mesênquima do prosencéfalo (UZ et al., 2010).

Os genes *homeobox* codificam proteínas de controle do desenvolvimento que controlam a formação do padrão embrionário e a ontogenia de tecidos específicos em organismos tão diversos quanto drosófilas, ratos, humanos, plantas, nematódeos e peixes-zebra (GEHRING; AFFOLTER; BURGLIN, 1994; WALTHER et al., 1991).

Em 1996, Gordon e colaboradores publicaram um estudo com o gene *CART-1*, que é o gene homólogo ao *ALX1* em ratos, o qual codifica uma proteína que é expressa na formação da cartilagem durante a diferenciação inicial de condrócitos. A sequência de aminoácidos identificada pelo estudo deste gene em humanos é 95% idêntica à dos ratos, indicando que a estrutura do *CART-1* é altamente conservada.

O *CART-1* mRNA foi detectado em altos níveis em células precursoras de condrócitos e condrócitos jovens. Estava presente em baixas quantidades em condrócitos maduros e não foi encontrado em tecidos não-cartilaginosos (ZHAO et al., 1993). Essa expressão restrita do *CART-1* em células precursoras de cartilagem sugerem que o *CART-1* pode estar envolvido no controle das células-tronco para a linhagem de condrócitos (GORDON et al., 1996) e com o desenvolvimento da cartilagem (ZHAO et



al., 1994). Como o desenvolvimento da cartilagem está intrinsecamente relacionado ao desenvolvimento ósseo, mutações neste gene podem levar a falhas no desenvolvimento ósseo.

No presente estudo, a mutação no gene *ALX1* foi identificada em um paciente com sintomas leves de OI. Na região da mutação houve uma deleção de sete aminoácidos. A proteína com esse segmento deletado sofre mudanças na porção anterior e posterior à mutação na estrutura secundária predita da proteína (Figuras 5 e 6). A figura 5 representa o esquema da estrutura secundária do *ALX1* selvagem e, a figura 6, representa o mesmo esquema, porém com a proteína mutada. Na posição 169 a 170 da proteína ocorre uma interrupção precoce na estrutura em  $\alpha$ -hélice. Na posição 177 a 180 ocorre uma interrupção prematura do segmento em  $\alpha$ -hélice alterando uma sequência de ligação proteica predita. Nas posições 300 a 307 e 310 a 313 há formação de uma  $\alpha$ -hélice não existente na proteína selvagem que invade parte da zona de ligação da proteína. Todas essas alterações podem influenciar na estrutura tridimensional e funcional da proteína *ALX1*.

Assim, infere-se que esta mutação é potencialmente causadora de OI.

#### 6.2.2. Gene *RYR1*

O gene *RYR1* codifica a proteína *Ryanodine receptor 1*, que funciona como um canal de cálcio que medeia a liberação de  $\text{Ca}^{2+}$  do retículo sarcoplásmico no citoplasma e, desse modo, desempenha um papel fundamental no desencadeamento da contração muscular (MONNIER et al., 2005; TILGEN et al., 2001).

As mutações neste gene estão associadas às doenças hipertermia maligna (MHS), doença do núcleo central do músculo (CCD) e doença multiminicora com oftalmoplegia externa (MMDO), todas relacionadas com o tecido muscular. A mais estudada e mais conhecida delas é a hipertermia maligna. A hipertermia maligna (MH) é uma miopatia potencialmente fatal desencadeada pela exposição a anestésicos voláteis e / ou relaxantes musculares despolarizantes. A susceptibilidade ao MH está associada principalmente a mutações dominantes no gene do receptor do tipo 1 de rianodina (*RYR1*). Estudos genéticos recentes mostraram que as variantes de *RYR1* são a causa mais comum de miopatias congênitas dominantes e recessivas - doença

central e multiminicora, desproporção congênita do tipo de fibra e miopatia centronuclear (MORONI et al., 1995).

A doença do núcleo central do músculo é uma miopatia congênita autossômica dominante, mas também existe uma forma autossômica recessiva grave. Observa-se variabilidade clínica e histológica. Indivíduos afetados tipicamente apresentam hipotonia e fraqueza muscular proximal na infância, levando ao atraso dos marcos motores. O curso clínico da doença geralmente é lento ou não progressivo na idade adulta e a gravidade dos sintomas pode variar de fraqueza muscular normal a significativa (ZHANG et al., 1993).

A doença multiminicora com oftalmoplegia externa é um transtorno neuromuscular heterogêneo. As características gerais incluem hipotonia neonatal, desenvolvimento motor retardado e fraqueza muscular generalizada e amiotrofia, que podem progredir lentamente ou permanecerem estáveis (JUNGBLUTH et al., 2005).

O *RYR1* também está envolvido no processo de formação do osso, ou de uma substância óssea, ou a conversão de tecido fibroso ou de cartilagem em osso, e na progressão do esqueleto desde a sua formação até o estado maduro (EBI, 2018). Também pode mediar a liberação de  $\text{Ca}^{2+}$  em neurônios e, desse modo, promover a sinalização prolongada de  $\text{Ca}^{2+}$  no cérebro. Necessário para o desenvolvimento embrionário normal das fibras musculares e do músculo esquelético. Necessário para a morfogênese cardíaca normal, desenvolvimento da pele e ossificação durante a embriogênese, por similaridade com o *RYR1* de ratos (UNIPROT, 2018).

Estudos demonstraram haver conexão entre os receptores de rianodina e o funcionamento de osteoblastos mediante a interferência no transporte de cátions  $\text{Ca}^{2+}$  nessas células. O Cálcio é um importante regulador da fixação, da mobilidade, da função endossômica (PIPER; LUZIO, 2004) e da sobrevivência celular (HAJNÓCZKY; DAVIES; MADESH, 2003). Adicionalmente, cálcio é usado por células de formação óssea e degradação óssea como um mediador secundário de hormônio e da ação da citosina. Essas ações incluem papéis na comunicação intercelular em grupos de osteoblastos que são conectados por “junções gap” (HENRIKSEN et al., 2006). Esses grupos de osteoblastos funcionam de forma coordenada na síntese e manutenção óssea e são coletivamente conhecidos como ósteons. Osteoblastos nesse grupo são conectados por junções gap que são capazes de propagar sinais nos grupos de células, incluindo ondas de cálcio (XIA; FERRIER, 1992). Canais de cálcio

adicionais podem estar envolvidos no acoplamento metabólico osteoblástico intracelular, incluindo o receptor de rianodina, receptores purinérgicos e IP3Rs (JØRGENSEN et al., 2003; SUN et al., 2002; WHYTE, 2017).

Assim, evidenciada a importante função do gene *RYR1* no metabolismo ósseo, infere-se que esta mutação é potencialmente causadora de OI.

## 7. CONSIDERAÇÕES ADICIONAIS

Apesar dos resultados deste trabalho terem identificado genes potencialmente patogênicos em OI, sabe-se que estes resultados não são conclusivos. Estudos funcionais em pesquisas futuras dos genes identificados são necessários para validar a hipótese de que os genes *ALX1* e *RYR1* estão relacionados com Osteogênese Imperfeita.

Além disso, destaca-se, neste trabalho, a utilização de estratégias de bioinformática com programas de fácil utilização como o Microsoft Office Excel, que torna mais acessível a pesquisa por profissionais das áreas Biológicas e de Saúde, já que a utilização da bioinformática, na maioria das vezes, requer conhecimentos específicos nas áreas de informática, sendo um campo de trabalho ocupado, principalmente, por cientistas e engenheiros da computação.

## 8. CONCLUSÕES

A partir de genes sequenciados por NGS em pacientes com Osteogênese Imperfeita que apresentavam diagnóstico molecular inconclusivo, foi possível, identificar genes candidatos potencialmente causadores de doença por meio de estratégias *in silico*. Os resultados sugerem que os genes *RYR1* e *ALX1* são potenciais alvos para estudos funcionais relacionados a OI. Estes resultados corroboram que os mecanismos de busca de genes por recursos de bioinformática são de grande importância para direcionar futuros estudos funcionais em busca de novos genes em doenças genéticas.

## 9. REFERÊNCIAS

- AUTON, A. et al. A global reference for human genetic variation. **Nature**, v. 526, n. 7571, p. 68–74, 30 set. 2015.
- BACON, S.; CROWLEY, R. Developments in rare bone diseases and mineral disorders. **Therapeutic Advances in Chronic Disease**, v. 9, n. 1, p. 51–60, 2018.
- BALJET, B. Aspects of the history of osteogenesis imperfecta (Vrolik's syndrome). **Annals of Anatomy**, v. 184, n. 1, p. 1–7, 2002.
- BALJET, B.; GROEN, G. J. Pathological blushing and sweating: Possible alternative autonomic pathways. **Clinical Neurology and Neurosurgery**, v. 86, n. 2, p. 122–124, jan. 1984.
- BERTOLA, D. R. et al. Vertical transmission of a frontonasal phenotype caused by a novel ALX4 mutation. **American Journal of Medical Genetics, Part A**, v. 161, n. 3, p. 600–604, 2013.
- BYERS, P. H. et al. Genetic evaluation of suspected osteogenesis imperfecta (OI). **Genetics in Medicine**, v. 8, n. 6, p. 383–388, 2006.
- CAI, R. L. Human CART1, a paired-class homeodomain protein, activates transcription through palindromic binding sites. **Biochemical and Biophysical Research Communications**, v. 250, n. 2, p. 305–311, 1998.
- CENTER FOR PEDIATRIC GENOMIC MEDICINE. Disponível em: <<https://www.ncbi.nlm.nih.gov/clinvar/submitters/505698/>>. Acesso em: 02 fev. 2018.
- EGL Genetic Diagnostics. Disponível em: <<https://www.ncbi.nlm.nih.gov/clinvar/submitters/500060/>>. Acesso em: 02 fev. 2018.
- EHLEN, H. W. et al. Inactivation of anoctamin-6/Tmem16f, a regulator of phosphatidylserine scrambling in osteoblasts, leads to decreased mineral deposition in skeletal tissues. **J Bone Miner Res**, v. 28, n. 2, p. 246–59, Feb 2013.
- GEHRING, W. J.; AFFOLTER, M.; BURGLIN, T. Homeodomain Proteins. **Annual Review of Biochemistry**, v. 63, n. 1, p. 487–526, 1994.
- GENEDX. Disponível em: <<https://www.ncbi.nlm.nih.gov/clinvar/submitters/26957/>>. Acesso em: 02 fev. 2018.

GONÇALVES, G. R.; MEYER, L. R.; SATO, M. T. Osteogênese Imperfeita: Relato De Caso. **Rev. Med. UFPR**, v. 4, n. 2, p. 87–92, 2017.

GORDON, D. F. et al. Human Cart-1: structural organization, chromosomal localization, and functional analysis of a cartilage-specific homeodomain cDNA. **DNA Cell Biol**, v. 15, n. 7, p. 531–541, 1996.

GRAY, P. H. K. A case of osteogenesis imperfecta, associated with dentinogenesis imperfecta, dating from antiquity. **Clinical Radiology**, v. 20, n. 1, p. 106–108, 1969.

HAJNÓCZKY, G.; DAVIES, E.; MADESH, M. Calcium signaling and apoptosis. **Biochemical and Biophysical Research Communications**, v. 304, n. 3, p. 445–454, 2003.

HENRIKSEN, Z. et al. The predominant mechanism of intercellular calcium wave propagation changes during long-term culture of human osteoblast-like cells. **Cell Calcium**, v. 39, n. 5, p. 435–444, 2006.

HUNG, N. N.; TAN, D.; HIEN, N. D. N. The Use of Kirschner Wire Intramedullary in Pediatric Femur Fractures with Osteogenesis Imperfecta. **International Journal of Orthopaedics**, v. 3, n. 6, p. 672–683, 2017.

ILLUMINA Clinical Services Laboratory. Disponível em: <<https://www.ncbi.nlm.nih.gov/clinvar/submitters/504895/>>. Acesso em: 02 fev. 2018.

INVITAE. Disponível em: <<https://www.ncbi.nlm.nih.gov/clinvar/submitters/500031/>>. Acesso em: 02 fev. 2018.

JIA, M. et al. Confirmation of the pathogenicity of a mutation p.G337C in the *COL1A2* gene associated with osteogenesis imperfecta. **Medicine (United States)**, v. 96, n. 39, 2017.

JØRGENSEN, N. R. et al. Activation of L-type Calcium Channels Is Required for Gap Junction-mediated Intercellular Calcium Signaling in Osteoblastic Cells. **Journal of Biological Chemistry**, v. 278, n. 6, p. 4082–4086, 2003.

JUNGBLUTH, H. et al. Minicore myopathy with ophthalmoplegia caused by mutations in the ryanodine receptor type 1 gene. **Neurology**, p. 65-77, 2005.

KANG, H.; ARYAL A.C., S.; MARINI, J. C. Osteogenesis imperfecta: new genes reveal novel mechanisms in bone dysplasia. **Translational Research**, v. 181, n. November,

p. 27–48, 2017.

KAYSERILI, H. et al. ALX4 dysfunction disrupts craniofacial and epidermal development. **Human Molecular Genetics**, v. 18, n. 22, p. 4357–4366, 2009.

LAMANDÉ, S. R.; BATEMAN, J. F. Collagen VI disorders: Insights on form and function in the extracellular matrix and beyond. **Matrix Biology**, 2018.

LIU, W. et al. Mutations in *AXIN2* cause colorectal cancer with defective mismatch repair by activating beta-catenin/TCF signalling. **Nature genetics**, v. 26, p. 146–147, 2000.

LOWENSTEIN, E. J. Osteogenesis imperfecta in a 3,000-year-old mummy. **Child's Nervous System**, v. 25, n. 5, p. 515–516, 2009.

MARDIS, E. R. The impact of next-generation sequencing technology on genetics. **Trends in Genetics**, v. 24, n. 3, p. 133–141, 2008.

MARINI, J. C. et al. Osteogenesis imperfecta. **Nature Reviews Disease Primers**, v. 3, p. 1–19, 2017.

MCGONNELL, I. M. et al. Evolution of the Alx homeobox gene family: Parallel retention and independent loss of the vertebrate Alx3 gene. **Evolution and Development**, v. 13, n. 4, p. 343–351, 2011.

METZKER, M. L. Sequencing technologies — the next generation. **Nature Reviews Genetics**, v. 11, n. 1, p. 31–46, 8 jan. 2010.

MONNIER, N. et al. Correlations between genotype and pharmacological, histological, functional, and clinical phenotypes in malignant hyperthermia susceptibility. **Human Mutation**, v. 26, n. 5, p. 413–425, 2005.

MORONI, I. et al. Ryanodine receptor gene point mutation and malignant hyperthermia susceptibility. **J Neurol**, v. 242, n. 3, p. 127–33, Feb 1995.

MOROZOVA, O.; MARRA, M. A. Applications of next-generation sequencing technologies in functional genomics. **Genomics**, v. 92, n. 5, p. 255–264, 2008.

O'RIORDAN, J. L. H.; BIJVOET, O. L. M. Rickets before the discovery of vitamin D. **BoneKEy Reports**, v. 3, 2014.

PIPER, R. C.; LUZIO, J. P. CUPpling calcium to lysosomal biogenesis. **Trends in Cell Biology**, v. 14, n. 9, p. 471–473, 2004.

SCHUSTER, S. C. Next-generation sequencing transforms today's biology. **Nature Methods**, v. 5, n. 1, p. 16–18, 2008.

SCHWARZ, J. M. et al. MutationTaster2: mutation prediction for the deep-sequencing age. **Nature Methods**, v. 11, n. 4, p. 361–362, 1 abr. 2014.

SILLENCE, D. O.; SENN, A.; DANKS, D. M. Genetic heterogeneity in osteogenesis imperfecta. **Journal of Medical Genetics**, v. 16, n. 2, p. 101–116, 1979.

ACMG: Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology, 2015. Disponível em: <<https://www.ncbi.nlm.nih.gov/pubmed/25741868>>. Acesso em: 02 fev. 2018.

SUN, L. et al. A novel mechanism for coupling cellular intermediary metabolism to cytosolic Ca<sup>2+</sup> signaling via CD38/ADP-ribosyl cyclase, a putative intracellular NAD<sup>+</sup> sensor. **FASEB journal : official publication of the Federation of American Societies for Experimental Biology**, v. 16, n. 3, p. 302–14, 2002.

TILGEN, N. et al. Identification of four novel mutations in the C-terminal membrane spanning domain of the ryanodine receptor 1: association with central core disease and alteration of calcium homeostasis. **Human molecular genetics**, v. 10, n. 25, p. 2879–2887, 2001.

TWIGG, S. R. F. et al. Frontorhiny, a Distinctive Presentation of Frontonasal Dysplasia Caused by Recessive Mutations in the ALX3 Homeobox Gene. **American Journal of Human Genetics**, v. 84, n. 5, p. 698–705, 2009.

UZ, E. et al. Disruption of *ALX1* Causes Extreme Microphthalmia and Severe Facial Clefting: Expanding the Spectrum of Autosomal-Recessive ALX-Related Frontonasal Dysplasia. **American Journal of Human Genetics**, v. 86, n. 5, p. 789–796, 2010.

WALTHER, C. et al. Pax: A murine multigene family of paired box-containing genes. **Genomics**, v. 11, n. 2, p. 424–434, 1991.

WEIL, U. H. Osteogenesis imperfecta: historical background. **Clinical orthopaedics and related research**, n. 159, p. 6–10, 1981.

WHYTE, M. P. Hypophosphatasia-aetiology, nosology, pathogenesis, diagnosis and treatment. **Nature Reviews Endocrinology**, v. 12, n. 4, p. 233–246, 2016.

WHYTE, M. P. Hypophosphatasia: An overview For 2017. **Bone**, v. 102, p. 15–25, 2017.

XIA, S. L.; FERRIER, J. Propagation of a calcium pulse between osteoblastic cells. **Biochemical and Biophysical Research Communications**, v. 186, n. 3, p. 1212–1219, 1992.

ZHANG, Y. et al. A mutation in the human ryanodine receptor gene associated with central core disease. **Nat Genet**, v. 5, n. 1, p. 46-50, Sep 1993

ZHAO, G. Q. et al. Cartilage Homeoprotein-1, a Homeoprotein Selectively Expressed in Chondrocytes. **Proceedings of the National Academy of Sciences of the United States of America**, v. 90, n. 18, p. 8633–8637, 1993.

ZHAO, G. Q. et al. The gene for the homeodomain-containing protein Cart-1 is expressed in cells that have a chondrogenic potential during embryonic development. **Mechanisms of Development**, v. 48, n. 3, p. 245–254, 1994.



## ANEXO I – AMOSTRA DA PLANILHA DO NGS COM OS PRINCIPAIS PARÂMETROS ANALISADOS.

	A	B	C	D	E	F	W	AF	AL	AP	AV	BA
1	Chr	Start	Alt	Func.refGene	Gene.refGene	ExonicFunc.refGene	avsift	AC	F9491-1	F9495-1	F9501-1	F9506-1
116	chr1	11852300	T	intronic	MTHFR	NA	NA	35	0/0:182,0:182:99	0/0:89,0:89:99	0/1:66,64:130:99	0/1:38,42:80:99
117	chr1	11852303	A	intronic	MTHFR	NA	NA	4	0/0:186,0:186:99	0/0:93,0:93:99	0/0:134,0:134:99	0/0:81,0:81:99:0
118	chr1	11852341	G	exonic	MTHFR	synonymous SNV	NA	1	0/0:227,1:226:99	0/0:135,0:135:	0/0:186,0:186:99	0/0:100,0:100:9
119	chr1	11852342	G	exonic	MTHFR	nonsynonymous SNV	0	1	0/0:228,0:228:99	0/0:138,1:138:	0/0:192,0:192:99	0/0:99,1:99:99:0
120	chr1	11854018	T	exonic	MTHFR	synonymous SNV	NA	4	0/0:249,0:249:99	0/0:193,0:193:	0/0:249,0:249:99	0/0:173,1:174:9
121	chr1	11854085	A	exonic	MTHFR	nonsynonymous SNV	0,01	1	0/0:245,0:246:99	0/0:201,1:208:	0/0:239,0:244:99	0/0:183,0:186:9
122	chr1	11854086	G	exonic	MTHFR	nonsynonymous SNV	0,85	1	0/0:246,2:246:99	0/0:210,0:210:	0/0:244,2:244:99	0/0:186,0:186:9
123	chr1	11854457	A	exonic	MTHFR	synonymous SNV	NA	168	1/1:0,246:246:99	1/1:0,246:245:	1/1:0,245:245:99	0/1:146,104:248
124	chr1	11854476	G	exonic	MTHFR	nonsynonymous SNV	0,12	43	0/0:245,4:247:99	0/0:244,1:244:	0/1:130,117:246:	0/1:145,104:248
125	chr1	11854755	A	intronic	MTHFR	NA	NA	109	0/1:100,107:207:	1/1:0,88:88:99	0/1:39,47:86:99:9	0/0:136,0:140:9
126	chr1	11854896	A	exonic	MTHFR	synonymous SNV	NA	8	0/0:228,0:228:99	0/0:99,1:100:9	0/0:138,0:138:99	0/1:58,80:137:9
127	chr1	11856378	A	exonic	MTHFR	nonsynonymous SNV	0	58	0/0:247,1:248:99	0/0:246,0:246:	0/0:249,0:249:99	0/0:250,0:250:9
128	chr1	11860342	T	exonic	MTHFR	synonymous SNV	NA	1	0/0:219,0:220:99	0/0:122,0:122:	0/0:189,4:192:99	0/0:104,0:104:9
129	chr1	11861345	A	exonic	MTHFR	synonymous SNV	NA	1	0/0:244,1:244:99	0/0:223,0:223:	0/0:246,0:246:99	0/0:242,0:242:9
130	chr1	11863057	A	exonic	MTHFR	synonymous SNV	NA	9	0/0:242,1:247:99	0/0:235,0:235:	0/0:245,0:245:99	0/1:99,127:224:
131	chr1	11866183	T	UTR5	CLCN6	NA	NA	10	0/0:153,0:153:99	0/0:79,1:80:99	0/0:129,0:129:99	0/0:98,0:98:99:0
132	chr1	11866184	G	UTR5	CLCN6	NA	NA	2	0/0:151,0:152:99	0/0:78,0:78:99	0/0:126,0:126:99	0/0:98,0:98:99:0
133	chr1	11994929	T	intronic	PLOD1	NA	NA	1	0/0:22,0:22:36:0,	0/0:21,0:21:33	0/0:10,0:11:15:0,	0/0:15,0:15:24:0
134	chr1	12008152	G	intronic	PLOD1	NA	NA	1	0/0:122,1:122:99	0/0:85,0:86:99	0/0:93,0:95:99:0,	0/0:102,0:102:9

Legenda:

- 1 - Localização da alteração genética
- 2 - Descrição da alteração genética
- 3 - Denominação do gene alterado
- 4 - Parâmetros da alteração genética
- 5 - Aspectos relativos à alteração genética nos pacientes avaliados

## CAMPOS QUE COMPÕEM A PLANILHA DO NGS

1 – **Localização da alteração genética:** Registros indicativos do cromossomo onde o gene alterado se encontra e posição inicial e final da alteração identificada.

Campos: **-Chr** – Informação relativa ao número do cromossomo onde o gene alterado está localizado.

**-Start** – Informação relativa à posição ordinal do nucleotídeo onde a alteração identificada inicia-se.

**-End** – Informação relativa à posição ordinal do nucleotídeo onde a alteração identificada encerra-se.

2 – **Descrição da alteração genética:** Registros indicativos dos nucleotídeos substituídos, inseridos ou deletados e à parte do gene onde a alteração foi identificada.

Campos: **-Ref** – Nucleotídeo ou conjunto de nucleotídeos que deveriam estar no local onde a alteração foi identificada, registrado nos campos “Start” e “End”. No caso de inserção, essa informação não é apresentada. No caso de deleção, o nucleotídeo ou conjunto de nucleotídeos que foram excluídos do DNA analisado.

**-Alt** – Nucleotídeo ou conjunto de nucleotídeos que substituíram aquele (s) informados no campo “Ref”. No caso de inserção, nucleotídeo ou conjunto de nucleotídeos inseridos na posição indicada no campo “Start”. No caso de deleção, essa informação não é apresentada.

**-Func.refGene** – Parte do gene onde a alteração foi identificada.

3 – **Denominação do gene alterado:** Registro indicativo do gene onde a alteração foi identificada.

Campo: **-Gene.refGene** – Denominação do gene ou, caso não se aplique, identificação da área do cromossomo onde a alteração foi identificada.

**4 – Parâmetros da alteração genética:** Registros indicativos das características, efeitos, tecidos afetados e outras informações relativas à alteração identificada.

Campos: Os campos de interesse para a pesquisa estão descritos na Metodologia, no item 4.5.1.

**5 – Aspectos relativos à alteração genética nos pacientes avaliados:** Registros que identificam os alelos afetados em cada paciente.